ЦЯНЬ СЮЭ-СЭНЬ

ТЕХНИЧЕСКАЯ КИБЕРНЕТИКА

ЦЯНЬ-СЮЭ-СЭНЬ

ТЕХНИЧЕСКАЯ КИБЕРНЕТИКА

Перевод с английского
М. З. ЛИТВИНА-СЕДОГО

Под редакцией
А. А. ФЕЛЬДБАУМА

ИЗДАТЕЛЬСТВО
ИНОСТРАННОЙ ЛИТЕРАТУРЫ
Москва, 1956

Цянь-Сюэ-Сэнь
ТЕХНИЧЕСКАЯ КИБЕРНЕТИКА

Редактор *А. В. ГЕРМОГЕНОВ*

Художник *М. В. Борисова-Мусатова*

Технический редактор *С. В. Клименко*

Сдано в производство 25/IV 1956 г.
Подписано к печати 4/VIII 1956 г.
Бумага 84×108 1/32=7,2 бум. л.
24,5 печ. л. Уч.-изд. л. 21,5.
Изд. № 1/2971
Цена 17 р. 05 к. Зак. № 311

ИЗДАТЕЛЬСТВО
ИНОСТРАННОЙ ЛИТЕРАТУРЫ
Москва, Ново-Алексеевская, 52.

16-я типография Главполиграфпрома
Министерства культуры СССР.
Москва, Трехпрудный пер., д. 9.

И * Л

Издательство

иностранной

литературы

*

ENGINEERING CYBERNETICS

H. S. TSIEN

Daniel and Florence Guggenheim Jet Propulsion Center
California Institute of Technology
Pasadena, California

McGraw-Hill Book Company

New York Toronto London

1954

ЦЯНЬ-СЮЭ-СЭНЬ

ТЕХНИЧЕСКАЯ КИБЕРНЕТИКА

Перевод с английского
М. З. ЛИТВИНА-СЕДОГО

Под редакцией
А. А. ФЕЛЬДБАУМА

ИЗДАТЕЛЬСТВО
ИНОСТРАННОЙ ЛИТЕРАТУРЫ
Москва, 1956

АННОТАЦИЯ

Книга Цяня «Техническая кибернетика» является первой выпускаемой на русском языке монографией по новому разделу науки — кибернетике. Кибернетика — наука о построении систем из механических и электрических компонентов для осуществления устойчивых целенаправленных действий. Книга рассчитана на широкие круги инженеров, интересующихся теорией этих вопросов, а также на научных работников и студентов старших курсов технических вузов.

Редакция литературы по математическим наукам
Заведующий редакцией профессор А. Г. КУРОШ

ПРЕДИСЛОВИЕ РЕДАКТОРА ПЕРЕВОДА

Последние годы характеризуются ярко выраженной тенденцией к созданию единой теоретической базы для огромного комплекса наук о передаче и преобразовании сигналов в целях связи и управления. Эту общую теоретическую базу называют по-разному: «теория систем управления», «общая теория информации», «общая теория связи», «техническая кибернетика», «общая техническая динамика» и т. д. Пути к этой новой, далеко еще не созревшей науке идут от теории связи, теории вычислительных устройств, теории автоматического регулирования и ряда других дисциплин. В книге Цяня «Техническая кибернетика» новая теория строится путем систематического обобщения задач автоматического регулирования. Поэтому она представляет наибольший интерес для специалистов по автоматическому регулированию.

Первые главы книги посвящены известным разделам теории регулирования и могут рассматриваться как введение в теорию. Однако уровень сложности и оригинальности изложения быстро возрастает вместе с последовательным обобщением задач. В последних главах кратко рассмотрены некоторые новые направления и поиски возможных новых общих путей развития автоматических систем.

В книге систематизирован большой материал, содержащий много интересных результатов, примеров и идей, причем использована обширная литература. Для удобства читателей в примечаниях к переводу добавлены ссылки на некоторые советские работы и учебники по автоматике.

А. А. Фельдбаум.

ПРЕДИСЛОВИЕ АВТОРА К РУССКОМУ ПЕРЕВОДУ

Выпуск русского издания моей книги «Техническая кибернетика», несомненно, сделает ее содержание более доступным большому и важному отряду ученых и инженеров современности—ученым и инженерам СССР. Я очень счастлив, что это случилось. Это произошло благодаря усилиям д-ра Фельдбаума и Издательства иностранной литературы. Д-р Фельдбаум добавил также библиографию работ советских ученых, а Издательство дало мне возможность внести некоторые исправления в русское издание. То и другое улучшает книгу. Поэтому я хочу воспользоваться этим случаем и выразить мою искреннюю благодарность д-ру Фельдбауму и Издательству.

Июль 1956 г. *Автор.*

ПРЕДИСЛОВИЕ АВТОРА

Знаменитый физик и математик А. М. Ампер ввел слово кибернетика (cibernétique) как термин, определяющий науку об управлении государством (часть II труда «Essai sur la philosophie des sciences», Paris, 1848). Грандиозная схема политических наук, построенная Ампером, не претворилась в жизнь и, возможно, не претворится никогда. Тем временем конфликты между правительствами, сопровождаемые применением силы, в значительной мере ускорили развитие другой ветви науки—науки о регулировании и управлении механическими и электрическими системами. И поэтому то обстоятельство, что Винер заимствовал это слово Ампера, окрестив им науку, столь важную для современной военной техники, носит, возможно, иронический оттенок. «Кибернетика» Винера (Cybernetics or control and communication in the animal and the mashine, New York, 1948) является наукой о построении систем из механических и электрических компонентов для осуществления устойчивых целенаправленных действий. Отличительной чертой

этой новой науки является полное отсутствие соображений, связанных с рассмотрением энергии, тепла и коэффициента полезного действия, столь важных в других естественных науках. По существу, главное внимание в кибернетике направлено на качественные стороны взаимодействий между различными компонентами системы и на поведение всей системы, обусловленное этими взаимодействиями.

Цель книги «Техническая кибернетика» состоит в изучении тех областей обширной науки, именуемой кибернетикой, которые имеют непосредственное техническое применение для разработки регулируемых, или управляемых систем. Эта книга содержит, конечно, те материалы по данной теме, которые обычно рассматриваются в руководствах по следящим системам. Но больший охват материала является лишь одним из различий между технической кибернетикой и прикладной теорией следящих систем. Более глубокое—и поэтому более важное—различие заключается в том, что техническая кибернетика представляет собой техническую *науку*, тогда как прикладная теория следящих систем представляет собой техническую *практику*.

Техническая наука имеет своею целью систематизировать принципы расчетов, применяемых в технической практике, чтобы построить единую научную дисциплину и тем самым выявить сходные черты в различных областях технической практики и показать мощь основных понятий. Короче говоря, в технической науке преобладает теоретическое исследование и очень часто применяются сложные методы математического анализа. Это отчетливо выступает уже при беглом просмотре содержания настоящей книги. Зато здесь почти не рассматриваются подробности конструктивного осуществления и проектирования компонентов систем, которые служат реальным приложением теории, и не приводится данных о конкретных устройствах.

Что же оправдывает это разделение теории и практики? В свете самого существования различных инженерных наук и их современного бурного развития такое оправдание едва ли является необходимым. Больше того, можно привести конкретный пример: механика жидкостей существует как инженерная наука отдельно от практики инженеров, работающих в области аэродинамики, гидравлики, метеорологии и других отраслей техники и пользующихся резуль-

татами исследований в области механики жидкостей в своей повседневной работе. В самом деле, без механики жидкостей понимание закономерностей сверхзвуковых потоков и их применение, несомненно, были бы сильно задержаны, чтобы не сказать больше. Поэтому образование технической кибернетики как технической науки оправдывается тем обстоятельством, что более широкий взгляд на предмет при систематизации рассмотрения часто может привести к новым и плодотворным методам подхода к старым задачам и открыть новые неожиданные перспективы. В настоящее время, когда существуют разнообразные применения техники автоматического регулирования и управления, вполне целесообразна попытка охватить все возможности, заложенные в этой новой науке, путем осмысленного обозрения всей ее области.

Поэтому рассуждения в области технической кибернетики должны с разумной степенью общности охватить все стороны науки, от которых можно ожидать технических приложений; в частности, не следует избегать какой-либо темы только по причине трудностей математического характера. Это тем более справедливо, что математические трудности всякого исследования обычно носят совершенно искусственный характер. При небольших изменениях в толковании материала изложение в общем случае можно упростить до уровня сложности, приемлемого для инженера-исследователя. Математический уровень данной книги отвечает познаниям студента, изучившего обычный курс математического анализа. Для понимания изложения необходимо знание теории интегрирования функций комплексной переменной, вариационного исчисления и обыкновенных дифференциальных уравнений. С другой стороны, мы не проводим строгих и изящных математических доказательств там, где достаточны эвристические соображения. Для специалиста-практика в области электроники трактовка, данная в книге, должна показаться чрезмерно «академичной», но для математика, интересующегося этой областью, она может показаться любительской. И если вся критика по адресу настоящей книги сведется только к этим замечаниям, то при всем их значении автор будет считать, что он не потерпел неудачи в выполнении поставленной перед собой задачи.

Глава I

ВВЕДЕНИЕ

Рассмотрим систему с одной степенью свободы, т. е. систему, физическое состояние которой можно охарактеризовать с помощью одной переменной y. Поведение такой системы описывается функциональной зависимостью переменной y от времени t. Для определения поведения системы или функции $y(t)$ необходимо знать структуру этой системы и свойства отдельных ее элементов. Эти сведения о системе вместе с использованием основных физических законов, переведенные на язык математики, доставляют уравнение для вычисления функции $y(t)$, которое может быть интегральным или интегродифференциальным, но весьма часто является дифференциальным уравнением, притом обыкновенным, так как здесь имеется только одна независимая переменная — время t.

Дифференциальное уравнение называется линейным, а система, описываемая дифференциальным уравнением, — линейной системой, если каждый член этого уравнения содержит самое большее только первую степень зависимой переменной y или ее производной по времени. В состав этого уравнения не должны входить более высокие степени y или смешанные произведения y и ее производных. В противном случае дифференциальное уравнение называется нелинейным, а система, описываемая таким уравнением, — нелинейной системой. Линейные системы можно, далее, разделить на системы с постоянными коэффициентами и системы с переменными коэффициентами. В системах с постоянными коэффициентами в качестве коэффициентов в членах дифференциального уравнения, описывающего систему, служат постоянные, не зависящие от времени. В системах с переменными коэффициентами коэффициенты являются функциями от t.

Приведенная здесь классификация типов дифференциальных уравнений оправдывается тем, что характер решения уравнения и тем самым поведение системы тесно связаны с типом описывающего ее дифференциального уравнения. Больше того, тип дифференциального уравнения определяет само существо вопросов, которые логически можно задавать относительно данной системы. Другими словами, тип дифференциального уравнения предопределяет и надлежащий путь решения технических задач, связанных с системой. Мы увидим это далее.

1. 1. Линейные системы с постоянными коэффициентами. Рассмотрим простейшую линейную систему — систему первого порядка. Это означает, что система описывается линейным дифференциальным уравнением первого порядка с постоянными коэффициентами. Если система свободная и не находится под действием «вынуждающих функций» («forced functions»), то ее дифференциальное уравнение можно записать в виде

$$\frac{dy}{dt}+ky=0, \tag{1.1}$$

где действительную величину k можно назвать упругой постоянной. Если y не меняется с течением времени, dy/dt равно нулю, и из уравнения (1.1) следует, что $y=0$. Таким образом, установившемуся, или равновесному, состоянию системы отвечает значение $y=0$.

Решением уравнения (1.1) служит функция

$$y=y_0e^{-kt}, \tag{1.2}$$

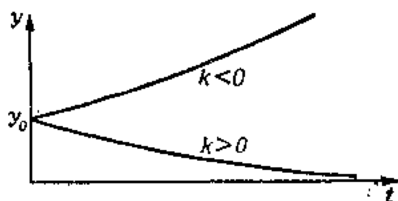
где y_0 — начальное значение y, т. е.

$$y(0)=y_0. \tag{1.3}$$

Следовательно, величина y_0 представляет собой начальное отклонение от равновесного состояния системы. График на фиг. 1 иллюстрирует поведение системы при $t>0$ как при положительном, так и при отрицательном k.

Из вида кривых следует, что при $k>0$ величина y убывает с течением времени и при неограниченном возрастании времени $y\to 0$. Поэтому при $k>0$ отклонение системы от равновесного состояния в конце концов прак-

тически обратится в нуль и, следовательно, систему можно назвать *устойчивой*. В случае $k < 0$ отклонение системы от положения равновесия увеличивается с течением времени, и в конце концов оно станет значительным, каким бы малым ни было начальное отклонение: система, однажды отклонившаяся от этого положения, никогда уже к нему не вернется. Такие системы, следовательно, *неустойчивы*.



Фиг. 1

Системы более высокого порядка описываются дифференциальными уравнениями, содержащими более высокие производные. Система n-го порядка описывается дифференциальным уравнением

$$\frac{d^n y}{dt^n} + a_{n-1}\frac{d^{n-1} y}{dt^{n-1}} + \ldots + a_0 y = 0. \qquad (1.4)$$

В физических системах коэффициенты $a_{n-1}, \ldots, a_0$ действительны. Решение уравнения (1.4) имеет вид

$$y = \sum_{i=1}^{n} y_0^{(i)} e^{\alpha_i t} \sin(\beta_i t + \varphi_i), \qquad (1.5)$$

где величины α_i, β_i — вещественные и зависят от коэффициентов $a_{n-1}, \ldots, a_0$, а величины φ_i представляют собой начальные фазы. Движение системы устойчиво лишь в том случае, когда все α_i отрицательны. Если хоть одна из этих величин положительна, то отклонение системы от состояния равновесия в конце концов будет возрастать неограниченно и, следовательно, система будет неустойчивой.

Примеры, разобранные выше, показывают, что при рассмотрении линейной системы с постоянными коэффициентами главным вопросом является вопрос об устойчивости. Едва ли требуется отмечать, что в технических расчетах обычно ставится цель обеспечения устойчивости системы. На вопрос же о том, будет ли система устойчивой или нет, можно ответить, коль скоро известны коэф-

фициенты дифференциального уравнения. В случае простой системы первого порядка, определяемой уравнением (1.1), существенен лишь знак коэффициента k.

1. 2. Линейные системы с переменными коэффициентами. Если рассматриваемая система содержит переменный параметр, то ее установившееся равновесное состояние можно изменять путем изменения этого параметра. Поэтому естественно ожидать, что коэффициенты линейного дифференциального уравнения, описывающего эту систему, также являются функциями от этого параметра. Например, аэродинамические силы, действующие на самолет, суть функции его скорости. Если скорость самолета меняется (в процессе ускорения или замедления его полета), происходит соответственное изменение аэродинамических сил, между тем как параметры самолета, характеризующие его инерцию, остаются практически неизменными. Вследствие этого основное дифференциальное уравнение, служащее для вычисления возмущенного полета самолета по отношению, например, к горизонтальному полету, будет уравнением с переменными коэффициентами.

Вернемся к простому примеру системы первого порядка, описываемой уравнением (1.1). Пусть упругая постоянная k является функцией скорости самолета; если самолет движется с постоянным ускорением a, то k представляет собой функцию величины $u = at$. Таким образом, дифференциальное уравнение принимает вид

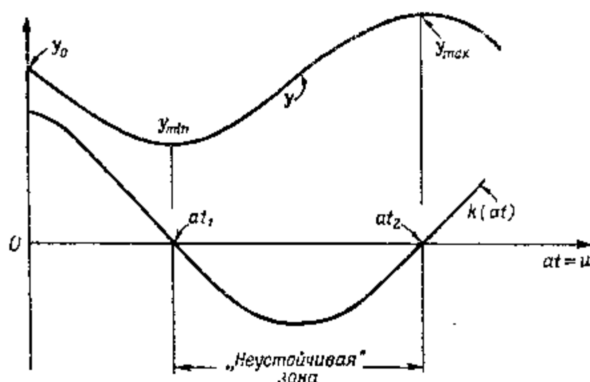
$$\frac{dy}{dt} + k(at)\,y = 0, \tag{1.6}$$

а его решение определяется интегралом

$$\ln\frac{y}{y_0} = -\frac{1}{a}\int_0^{at} k(\xi)\,d\xi, \tag{1.7}$$

где y_0 — начальное возмущение. Если k всегда остается положительным, то $\ln(y/y_0)$ всегда будет отрицательным и с течением времени этот логарифм будет возрастать по абсолютной величине, оставаясь отрицательным. Следовательно, y всегда будет меньше y_0 и в конце концов (практически) обратится в нуль. Таким образом, в этом

случае система устойчива. Если k всегда отрицательно, то $\ln(y/y_0)$ остается положительным и с течением времени возрастает. Следовательно, y со временем сделается весьма большим, даже если начальное отклонение y_0 очень мало. В этом случае система неустойчива. Отметим, что указанные здесь характеристики линейной системы с переменными коэффициентами, остающиеся всегда положительными или отрицательными, весьма похожи на соответственные характеристики систем с постоянными коэффициентами.



Фиг. 2

Представляет интерес тот случай, когда k принимает как положительные, так и отрицательные значения. Пусть $k(at)$ сначала положительна, затем отрицательна, а потом опять положительна. Обозначим первый нуль функции k через $u_1 = at_1$, а второй — через $u_2 = at_2$. Тогда, согласно предыдущему, система должна быть неустойчивой в зоне изменения скорости от u_1 до u_2 (фиг. 2). Пусть $y_{\min}$ — минимальное значение y, а $y_{\max}$ — его максимальное значение. Тогда равенство (1.7) дает

$$\ln \frac{y_{\min}}{y_0} = -\frac{1}{a} \int_0^{u_1} k(\xi)\, d\xi \tag{1.8}$$

и

$$\ln \frac{y_{\max}}{y_0} = -\frac{1}{a} \int_0^{u_2} k(\xi)\, d\xi. \tag{1.9}$$

В технической задаче первостепенное значение имеет вопрос: как велико y_{max}? Не становится ли y_{max} настолько большим, что система уже не сможет отвечать заданным требованиям? Чтобы ответить на этот вопрос, необходимо знать, помимо функционального выражения зависимости k от u, еще два параметра: величину a ускорения и величину y_0 начального отклонения. При любом заданном a величина y_{max} пропорциональна y_0. Но еще более важное свойство заключается в том, что при всяком заданном начальном отклонении максимальную величину y_{max} отклонения системы от равновесного состояния можно в значительной мере снизить путем увеличения ускорения a, как это следует из равенства (1.9). Последнее свойство означает, что нежелательные явления в системе можно сделать менее заметными при возможно более быстром прохождении «неустойчивой» зоны.

Таким образом, в более общем случае линейной системы с переменными коэффициентами простой вопрос об устойчивости не имеет определенного смысла. Большее значение имеет вопрос о критерии, выполнение которого обусловливало бы удовлетворительную работу системы при заданных начальных возмущениях и других условиях. В рассмотренной системе первого порядка таким критерием качества работы системы служит величина y_{max}; заданное возмущение равно y_0, а другие условия сводятся к заданию ускорения a. Таким образом, переход от систем с постоянными коэффициентами к системам с переменными коэффициентами приводит к значительному изменению самого характера задачи[1]).

1.3. Нелинейные системы. Если в простой системе первого порядка, описываемой дифференциальным уравнением (1.1), упругая постоянная k является функцией от самого отклонения y системы, то это уравнение имеет

---

[1]) С этим положением нельзя полностью согласиться, так как и в случае систем с постоянными коэффициентами большой интерес представляет не только асимптотическая устойчивость, но и максимум отклонения, время затухания переходного процесса и другие условия, обычно называемые условиями качества регулирования. — *Прим. ред.*

вид

$$\frac{dy}{dt}+f(y)=0. \qquad (1.10)$$

где $f(y)=k(y)y$. Очевидно, что это дифференциальное уравнение — нелинейное. Система, описываемая уравнением (1.10), представляет собой, таким образом, простейший пример нелинейной системы. Интегрируя это уравнение, можно найти его решение $y(t)$ из соотношения

$$t=-\int_{y_0}^{y}\frac{d\eta}{f(\eta)}, \qquad (1.11)$$

где, как и выше, y_0 — начальное отклонение.

С другой стороны, повторно дифференцируя обе части уравнения (1.10), получаем

$$\left.\begin{aligned} \frac{d^2y}{dt^2}+\frac{df}{dy}\frac{dy}{dt}&=0,\\ \frac{d^3y}{dt^3}+\frac{d^2f}{dy^2}\left(\frac{dy}{dt}\right)^2+\frac{df}{dy}\frac{d^2y}{dt^2}&=0,\\ \ldots\ldots\ldots\ldots\ldots\ldots& \end{aligned}\right\} \qquad (1.12)$$

Таким образом, если число y_1 является нулем функции $f(y)$ и $f(y)$ регулярна при $y=y_1$, т. е. если производные всех порядков от функции $f(y)$ принимают при $y=y_1$ конечные значения, то из уравнений (1.10) и (1.12) вытекают равенства

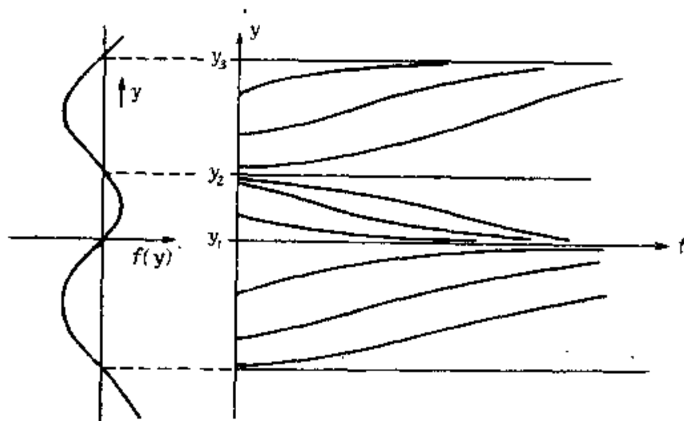
$$\frac{dy}{dt}=\frac{d^2y}{dt^2}=\frac{d^3y}{dt^3}=\ldots=0 \text{ при } y=y_1. \qquad (1.13)$$

Эти равенства означают, что переменная y асимптотически приближается к значению y_1. Действительно, если $y_0>y_1$ и $f(y_0)>0$, то y достигает y_1 лишь при $t\to\infty$. Если же $y_0<y_1$, то $f(y_0)<0$ и y опять-таки достигает y_1 при $t\to\infty$[1]). Таков же характер приближе-

[1]) Мотивировка автора неясна, однако утверждение верно. Так, если $f(\eta)=(\eta-y_1)^k[C_k+C_{k+1}(\eta-y_1)+\ldots]$, где $C_k\neq 0$, $k\geqslant 1$, то из равенства (1.11) видно, что при $y\to y_1$ t ведет себя в основном как интеграл $-\int_{y_0}^{y}\frac{d\eta}{(\eta-y_1)^k}$, т. е. $t\to\infty$. — *Прим. ред*

ния y и к остальным нулям функции $f(y)$ (фиг. 3)[1]).

Если величина начального отклонения совпадает с одним из нулей функции $f(y)$, то это значение y будет с течением времени сохраняться неизменным. Таким образом, значения y, представляющие собой нули функции $f(y)$, соответствуют положениям равновесия. При $df/dy > 0$ в окрестности некоторого нуля, например y_1 на фиг. 3, малые отклонения системы от этого равновесного положения с течением времени затухнут и система в конечном



Фиг. 3

счете вернется к исходному состоянию; в этом случае можно сказать, что система обладает устойчивостью по отношению к малым возмущениям в y_1. Если же $df/dy < 0$ в окрестности некоторого нуля, например y_2 на фиг. 3, то малейшее отклонение от соответствующего положения равновесия заставит систему перемещаться в направлении одного из двух соседних равновесных положений, определяемых координатами y_1 или y_3 на этой фигуре. Таким образом, состояние равновесия системы, определяемое координатой y_2, неустойчиво.

Мы убедились, насколько сложно ведет себя даже очень простая нелинейная система, описываемая уравне-

---

[1]) См. также [1], гл. IV, § 5 и 6. — *Прим. перев.*

нием (1.10). Система может обладать в зависимости от данного состояния как устойчивостью, так и неустойчивостью. Следовательно, для таких систем совершенно бессмысленно ставить общий вопрос об устойчивости [1]); вместо этого каждую конкретную задачу надо рассматривать в отдельности.

1.4. Инженерное приближение задачи. Почти очевидно, что любая физическая система при ее тщательном исследовании всегда окажется нелинейной. Говоря о системе, как о линейной, мы имеем в виду лишь то обстоятельство, что эту систему можно с достаточной точностью аппроксимировать линейной системой. Понятие же достаточной точности определяется в следующем смысле: отклонение системы от ее линейного приближения настолько мало, что в рассматриваемой задаче влияние ее нелинейных свойств не является существенным. Таким образом, для отнесения данной системы к классу линейных или нелинейных систем необходимо исходить из точно сформулированных условий, при которых рассматривается эта система. Общего же абсолютного критерия не существует [2]).

Такое же заключение можно высказать по отношению к задаче о классификации линейных систем на системы с постоянными коэффициентами и системы с переменными коэффициентами. Обратимся к простым примерам систем, описываемых уравнениями (1.1) и (1.6). Если ускорение a очень мало, т. е. скорость почти постоянна, то, как следует из (1.8), величина y_{min} во много

[1]) Как известно, для самых общих случаев понятию устойчивости может быть дан совершенно определенный смысл. Такой смысл имеет устойчивость по Ляпунову и другие родственные и близкие этому понятия (см. [2—4]). — *Прим. ред.*

[2]) Например, если нас интересует асимптотическая устойчивость некоторого установившегося состояния нелинейной системы по отношению к достаточно малым начальным возмущениям, то, по теореме А. М. Ляпунова [2—4] об устойчивости по первому приближению, исходную систему можно при определенных условиях аппроксимировать линейной. Если же ставится задача о разыскании автоколебаний, то нелинейностью пренебрегать нельзя, как бы мала она ни была, ибо такое пренебрежение уничтожит физическую (и математическую) основу, обусловливающую возможность установления автоколебаний. — *Прим. перев.*

раз меньше величины y_0 начального отклонения и переменная y примет значение $y_{\min}$ при очень большом значении t. Поэтому поведение системы в течение некоторого конечного промежутка времени весьма схоже с поведением системы, описываемой уравнением (1.1) при положительном k. Следовательно, эту систему с переменными коэффициентами при определенных условиях можно с достаточной точностью аппроксимировать системой с постоянными коэффициентами.

Едва ли следует отдельно отмечать, что линейные системы с постоянными коэффициентами оказываются наиболее легкими для исследования. К счастью, под категорию линейных систем подпадает очень большое число реальных систем, встречающихся в технике, при соответствующих условиях «технической аппроксимации». Именно по этой причине наибольшее развитие получила та область теории устойчивости и автоматического регулирования, которая ограничивается линейными задачами. В действительности современная теория следящих систем охватывает почти исключительно линейные задачи [1]). Поэтому мы начнем с рассмотрения линейных систем с постоянными коэффициентами.

---

[1]) Следует отметить, что за последние годы число исследований по нелинейным системам регулирования и нелинейным следящим системам (не говоря об общей теории устойчивости и нелинейных колебаний) значительно возросло. См., например, [5—9]. —*Прим. перев.*

Глава II

МЕТОД ПРЕОБРАЗОВАНИЯ ЛАПЛАСА

Для нахождения решений линейных дифференциальных уравнений с постоянными коэффициентами, в которых независимым переменным служит время t, особенно полезен метод, основанный на применении преобразования Лапласа. Разумеется, эту задачу можно решить и многими другими способами; однако метод преобразования Лапласа в большей мере подходит для инженеров-исследователей благодаря тому, что упомянутый метод сводит все проблемы к единой основе. Процесс нахождения решения стандартизируется, что делает возможным общий подход. Теория и практические методы применения преобразования Лапласа изложены во многих руководствах [1] [2]), и в настоящей главе мы не будем этим заниматься. Наша цель здесь состоит скорее в том, чтобы дать удобный для справок обзор соответствующих результатов, используемых в дальнейших главах. Для выяснения подробностей, а также для ознакомления с доказательствами читателю следует обратиться к руководствам, указанным в примечаниях 1) и 2).

2.1. Преобразование Лапласа и формула обращения. Пусть $y(t)$ — функция независимого переменного t (времени), определенная при $t > 0$. Тогда изображение $Y(s)$.

---

[1]) См., например, Х. Карслоу и Д. Егер, Операционные методы в прикладной математике, М., 1948; Churchill R. V., Modern operational methods in engineering, New York, 1944.

Более подробно теория преобразования Лапласа рассмотрена в книгах: Doetsch. G., Theorie und Anwendung der Laplace-Transformation, Berlin, 1937; Widder D. V., The Laplace transform, Princeton, 1946.

[2]) См. также [9—15]. — *Прим. перев.*

по Лапласу функции $y(t)$ определяется равенством[1])

$$Y(s)=\int_0^\infty e^{-st}y(t)\,dt, \tag{2.1}$$

где s — комплексная переменная, обладающая положительной вещественной частью, $\operatorname{Re} s>0$. При других значениях s функция $Y(s)$ определяется с помощью аналитического продолжения. Размерность $Y(s)$ равна размерности y, умноженной на время.

Если изображение $Y(s)$ известно, то оригинал, т. е. функцию, изображением которой по Лапласу служит $Y(s)$, можно во всех случаях найти с помощью формулы обращения

$$y(t)=\frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty} e^{st}Y(s)\,ds, \tag{2.2}$$

где γ — постоянная, превосходящая вещественные части всех особых точек изображения $Y(s)$. Фактическое определение функции $y(t)$ можно выполнить путем соответствующей деформации контура интегрирования, отвечающей характеру изображения $Y(s)$.

2.2. Применение преобразования Лапласа к линейным уравнениям с постоянными коэффициентами. Так как преобразование Лапласа вводится как операция над функцией, определенной при $t>0$, то этот метод оказывается специально приспособленным к решению задач с начальными условиями: задано начальное состояние системы и, при $t>0$, дана вынуждающая функция, действующая на систему; требуется найти «движение» системы при $t>0$. Рассмотрим систему n-го порядка, описываемую линейным дифференциальным уравнением с коэффициентами a_n, a_{n-1}, ..., a_0 при соответствующих производных и вынуждающей функцией $x(t)$, т. е. уравнением

$$a_n\frac{d^ny}{dt^n}+a_{n-1}\frac{d^{n-1}y}{dt^{n-1}}+\ldots+a_0y=x(t). \tag{2.3}$$

---

[1]) На протяжении всей книги прописные буквы относятся к изображениям переменных, обозначенных соответствующими строчными буквами.

Начальные значения производных задаются обычно в следующем виде:

$$\left.\begin{aligned} &\left(\frac{d^{n-1}y}{dt^{n-1}}\right)_{t=0} = y_0^{(n-1)}, \\ &\ldots\ldots\ldots\ldots \\ &(y)_{t=0} = y_0. \end{aligned}\right\} \tag{2.4}$$

Дифференциальное уравнение (2.3) вместе с начальными условиями (2.4) определяет единственным образом поведение системы при $t > 0$.

Для решения задачи о движении системы с помощью преобразования Лапласа умножим обе части уравнения (2.3) на e^{-st} и проинтегрируем от $t = 0$ до $t = \infty$. Тогда

$$\int_0^\infty e^{-st} y(t)\, dt = Y(s). \tag{2.1a}$$

Изображения для производных $\frac{dy}{dt}, \frac{d^2y}{dt^2}, \ldots, \frac{d^ny}{dt^n}$ получаем, выполняя интегрирование по частям:

$$\begin{aligned} &\int_0^\infty e^{-st} \frac{dy}{dt}\, dt = -y_0 + s\int_0^\infty e^{-st} y(t)\, dt = -y_0 + sY(s), \\ &\int_0^\infty e^{-st} \frac{d^2y}{dt^2}\, dt = -y_0^{(1)} - sy_0 + s^2 Y(s) \\ &\qquad\ldots\ldots\ldots \end{aligned} \tag{2.5}$$

и

$$\int_0^\infty e^{-st} \frac{d^ny}{dt^n}\, dt = -y_0^{(n-1)} - sy_0^{(n-2)} - \ldots - s^{n-1}y_0 + s^n Y(s).$$

Поэтому, обозначая через $X(s)$ изображение по Лапласу вынуждающей функции $x(t)$

$$X(s) = \int_0^\infty e^{-st} x(t)\, dt, \tag{2.6}$$

мы можем записать изображение уравнения (2.3) с учетом начальных условий (2.4) в виде

$$(a_n s^n + a_{n-1}s^{n-1} + \ldots + a_1 s + a_0)\, Y(s) = a_n y_0 s^{n-1} + \\ + (a_n y_0^{(1)} + a_{n-1}y_0)\, s^{n-2} + (a_n y_0^{(2)} + a_{n-1}y_0^{(1)} + a_{n-2}y_0)\, s^{n-3} + \\ + \ldots + (a_n y_0^{(n-1)} + a_{n-1}y_0^{(n-2)} + \ldots + a_1 y_0) + X(s). \quad (2.7)$$

Обозначив через $D(s)$ и $N_0(s)$ многочлены

$$D(s) = a_n s^n + a_{n-1}s^{n-1} + \ldots + a_1 s + a_0 \quad (2.8)$$

и

$$N_0(s) = a_n y_0 s^{n-1} + (a_n y_0^{(1)} + a_{n-1}y_0)\, s^{n-2} + \ldots + \\ + (a_n y_0^{(n-1)} + a_{n-1}y_0^{(n-2)} + \ldots + a_1 y_0), \quad (2.9)$$

напишем решение операционного уравнения (2.7) в виде

$$Y(s) = \frac{N_0(s)}{D(s)} + \frac{X(s)}{D(s)}. \quad (2.10)$$

Заметим, что решение (2.10) зависит от начальных условий через посредство многочлена $N_0(s)$, входящего в первый член правой части (2.10). Степень многочлена $N_0(s)$ не превышает $n-1$, и, таким образом, она всегда меньше степени $D(s)$. $N_0(s)$ обращается в нуль при обращении в нуль всех начальных значений, задаваемых равенствами (2.4). В этом случае изображение $Y(s)$ определяется только вторым членом правой части (2.10), зависящим от вынуждающей функции. Поэтому первый член, $N_0(s)/D(s)$, можно назвать дополнительной функцией, а второй член, $X(s)/D(s)$, — частным интегралом. Фактическое решение $y(t)$ можно найти по его изображению $Y(s)$, применяя формулу обращения (2.2).

2.3. «Словарь» преобразования Лапласа. Вынуждающая функция $x(t)$ часто имеет такой вид, при котором ее изображение $X(s)$ является отношением двух многочленов относительно s. Тогда изображение $Y(s)$ общего решения, определяемое соотношением (2.10), также выражается в виде отношения двух многочленов относительно s, и его можно разложить на элементарные дроби.

Оригиналы, отвечающие каждой такой дроби, можно получить с помощью формулы обращения или же, что более удобно, с помощью заранее подготовленного «словаря», содержащего перечень некоторых употребительных функций от t и их изображений по Лапласу. Такой очень краткий перечень приведен в табл. 1.

Таблица 1

«СЛОВАРЬ» ПРЕОБРАЗОВАНИЯ ЛАПЛАСА

$Y(s)$	$y(t)$
$1/s$	1
$1/s^n$	$t^{n-1}/\Gamma(n)$
$1/s-a$	e^{at}
$a/(s^2+a^2)$	$\sin at$
$s/(s^2+a^2)$	$\cos at$
$a/(s^2-a^2)$	$\operatorname{sh} at$
$s/(s^2-a^2)$	$\operatorname{ch} at$
$s/(s^2+a^2)^2$	$\frac{t}{2a}\sin at$
$1/(s^2+a^2)^2$	$\frac{1}{2a^3}(\sin at - at\cos at)$

2.4. Синусоидальная вынуждающая функция. Рациональную дробь $N_0(s)/D(s)$ можно разложить на элементарные дроби. Если все корни многочлена $D(s)$ различны, то, обозначив их через $s_1, s_2, \ldots, s_n$, мы можем написать

$$\frac{N_0(s)}{D(s)} = \sum_{r=1}^{n} \frac{N_0(s_r)}{D'(s_r)} \frac{1}{s-s_r},$$

где $D'(s)$ — производная от $D(s)$ по s. «Интерпретируя» каждый член суммы в правой части последнего равенства с помощью нашего «словаря» изображений и оригиналов, выпишем слагаемое $y_c(t)$ полного решения, обусловленное начальными условиями, т. е. дополнительную функцию[1]),

[1]) $y_c(t)$ — решение однородного уравнения, соответствующего данному неоднородному, при заданных начальных условиях. — *Прим. перев.*

в виде

$$y_c(t) = \sum_{r=1}^{n} \frac{N_0(s_r)}{D'(s_r)} e^{s_r t}. \tag{2.11}$$

В общем случае корни s_r многочлена $D(s)$ суть комплексные числа. В физических системах коэффициенты $a_0, \ldots, a_n$ многочлена $D(s)$ (2.8) вещественны; поэтому s_r могут быть лишь попарно сопряженными комплексными числами. Если вещественные части всех чисел s_r отрицательны, то функция $y_c(t)$ убывает с течением времени по экспоненциальному закону и стремится к нулю при неограниченном возрастании t. Таким образом, в этом случае система устойчива.

Если вынуждающая функция *синусоидальная*, то ее можно записать в виде

$$x(t) = x_m e^{i\omega t}, \tag{2.12}$$

где x_m — амплитуда, а ω — круговая частота; обращаясь к таблице изображений и оригиналов (стр. 23), мы видим, что

$$X(s) = x_m \frac{1}{s - i\omega}.$$

Поэтому второе слагаемое в правой части равенства (2.10) принимает вид

$$\frac{x_m}{(s - i\omega) D(s)}.$$

Последнее выражение можно обобщить на случай системы уравнений, добавив в числителе множитель в виде некоторого другого многочлена $N(s)$, степень которого меньше n. Тогда изображение $Y_i(s)$ частного интеграла можно написать в виде

$$Y_i(s) = F(s) X(s) = \frac{N(s)}{D(s)} X(s) = \frac{x_m N(s)}{(s - i\omega) D(s)}. \tag{2.13}$$

При $N(s) \equiv 1$ мы возвращаемся к более простому случаю, отвечающему соотношению (2.10). По отношению к изображению $Y(s)$ можно опять применить метод разложения на элементарные дроби. Но многочлен в знаменателе есть теперь $(s - i\omega) D(s)$, а его корни суть s_1,

$s_2, \ldots, s_n$ и $i\omega$. Таким образом,

$$Y_i(s) = \left[\frac{N(i\omega)}{D(i\omega)} \frac{1}{(s - i\omega)} + \sum_{r=1}^{n} \frac{N(s_r)}{(s_r - i\omega) D'(s_r)} \frac{1}{s - s_r} \right] x_m. \quad (2.14)$$

Поэтому частный интеграл $y_i(t)$, обусловленный наличием синусоидальной вынуждающей функции (2.12), имеет вид

$$y_i(t) = x_m \left[\frac{N(i\omega)}{D(i\omega)} e^{i\omega t} + \sum_{r=1}^{n} \frac{N(s_r)}{(s_r - i\omega) D'(s_r)} e^{s_r t} \right]. \quad (2.15)$$

В устойчивых системах вещественные части всех чисел s_r отрицательны, и потому второе слагаемое в правой части равенства (2.15) при $t \to \infty$ обращается в нуль. Первое же слагаемое представляет собой установившееся решение, и, таким образом, отношение установившегося решения к вынуждающей функции определяется простым выражением:

$$\frac{[y_i(t)]_{\text{уст.}}}{x(t)} = \frac{N(i\omega)}{D(i\omega)} = F(i\omega). \quad (2.16)$$

Это равенство дает прямой путь вычисления установившегося решения при синусоидальной вынуждающей функции.

При уменьшении частоты ω вынуждающей функции до нуля сама вынуждающая функция сводится к постоянной, т. е. к величине, не изменяющейся с течением времени. Соотношение (2.16) показывает тогда, что $F(0)$ равно отношению y к x при x постоянном. Это равенство определяет физический смысл величины $F(s)$ при $s = 0$. В дальнейших рассуждениях мы будем часто прибегать к этому физическому истолкованию.

2.5. Реакция системы на единичный импульс. Вынуждающая функция $x(t)$ не обязательно является непрерывной. Она может иметь форму единичного импульса, приложенного к системе в момент времени $t = 0$, т. е.

$$x(t) = 0 \quad \text{при} \quad t \neq 0,$$

$$x(t) \to \infty \quad \text{при} \quad t \to 0$$

и

$$\int_0^\infty x(t)\,dt = 1.$$

Изображение по Лапласу $X(s)$ единичного импульса, определенного указанным образом, просто равно единице. Изображение реакции системы, обобщенной согласно (2.13), на единичный импульс выражается простым равенством:

$$Y_i(s) = \frac{N(s)}{D(s)} \cdot 1 = F(s). \tag{2.17}$$

Решение $y(t)$, соответствующее этой единичной импульсной функции, обычно обозначается через $h(t)$. На основании формулы обращения (2.2)

$$h(t) = \frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} e^{st} F(s)\,ds. \tag{2.18}$$

Когда система устойчива, вещественные части всех корней s_r характеристического уравнения системы $D(s) = 0$ отрицательны. Поэтому все особые точки изображения $F(s)$ лежат по левую сторону от мнимой оси на плоскости комплексной переменной s, а в таком случае эту ось можно выбрать в качестве контура интегрирования при подсчете $h(t)$, т. е. в равенстве (2.18) можно принять $\gamma = 0$.

Глава III

ВХОДНАЯ, ВЫХОДНАЯ И ПЕРЕДАТОЧНАЯ ФУНКЦИИ

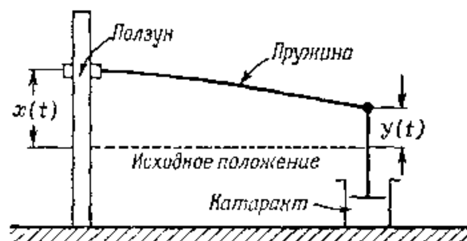
В предыдущей главе мы видели, что при применении преобразования Лапласа задача изучения поведения линейной системы с постоянными коэффициентами оказывается существенно связанной с исследованием многочлена $D(s)$ [равенство (2.8)], образуемого с помощью коэффициентов дифференциального уравнения системы. В более общем случае, когда физическая система описывается системой дифференциальных уравнений, при нулевых начальных значениях всех переменных, определяющих состояние системы, поведение системы полностью определяется отношением $N(s)/D(s)$ двух многочленов. Обозначим это отношение через $F(s)$. Пусть $X(s)$ — лапласово изображение вынуждающей функции, а $Y_i(s)$ — лапласово изображение частного интеграла. Тогда на основании (2.13) получим

$$Y_i(s) = F(s)X(s). \tag{3.1}$$

Это уравнение можно рассматривать как операторное уравнение; при действии $F(s)$ на $X(s)$ последняя функция преобразуется в $Y_i(s)$, или, иначе, $F(s)$, действуя на $X(s)$, преобразует $X(s)$ в $Y_i(s)$. По этой причине $F(s)$ называют *преобразующей* или *передаточной функцией*. $X(s)$ представляет собой лапласово изображение *входной функции* $x(t)$, а $Y_i(s)$ — лапласово изображение *выходной функции* $y_i(t)$. Чтобы подчеркнуть то обстоятельство, что функция $y_i(t)$ определяет только частный интеграл, без учета дополнительной функции, обусловливаемой влиянием начальных условий в системе, $y_i(t)$ называют *выходной функцией* (или просто *выходом*) *от входного воздействия*. В соответствии с этим дополнительную функцию $y_c(t)$ называют *выходом от начальных условий*.

Преимущество метода преобразования Лапласа заключается в сведении задачи, характеризуемой дифференциальными уравнениями, к задаче, связанной с алгебраическими действиями. При этом обратный переход от $Y(s)$ к $y(t)$ оказывается необходимым лишь в редких случаях, так как свойства функции $y(t)$ полностью определяются ее изображением $Y(s)$. Таким образом, появляется возможность выразить заданные технические требования к функции $y(t)$ через посредство определенных требований к функции $Y(s)$ или, при заданном входном воздействии, через посредство требований к передаточной функции $F(s)$. В технике следящих систем основным является метод исследования и проектирования с помощью передаточной функции. В настоящей главе мы проиллюстрируем этот метод на нескольких примерах.

3.1. Системы первого порядка. В качестве первого примера рассмотрим движение скользящей пружины, один



Фиг. 4

конец которой скреплен с поршнем катаракта, а другой скользит вдоль направляющей (фиг. 4). Обозначим через $y(t)$ координату конца пружины, скрепленного с поршнем, а через $x(t)$ — координату ее скользящего конца. Благодаря влиянию катаракта координата $y(t)$ не равна координате $x(t)$, а именно $y(t)$ отстает от $x(t)$[1].

Задача состоит в исследовании изменения выходной переменной $y(t)$, когда концу пружины, связанному с направ-

[1] Автор пренебрегает инерцией пружины, что он п оговаривает ниже. — *Прим. перев.*

ляющей, задается определенное перемещение. Таким образом, переменная $x(t)$ является входной.

Обозначим через k упругую постоянную пружины, а через c — коэффициент демпфирования, производимого катарактом. Тогда, считая движение настолько медленным, что силами инерции можно пренебречь [1]), выпишем условие равновесия сил

$$c\frac{dy}{dt}+k(y-x)=0.$$

Введя постоянную времени τ_1, определяемую отношением

$$\tau_1=\frac{c}{k}, \tag{3.2}$$

перепишем уравнение движения в виде

$$\tau_1\frac{dy}{dt}+y=x. \tag{3.3}$$

Начальное условие задачи выражается простым равенством

$$y(0)=y_0. \tag{3.4}$$

Умножив обе части уравнения (3.3) на e^{-st} и проинтегрировав от $t=0$ до $t=\infty$, получим уравнение в изображениях

$$(\tau_1 s+1)Y(s)=X(s)+\tau_1 y_0,$$

откуда

$$Y(s)=\frac{X(s)}{\tau_1 s+1}+\frac{\tau_1 y_0}{\tau_1 s+1}. \tag{3.5}$$

Таким образом, изображение выходной функции от входного воздействия определено функцией

$$Y_i(s)=\frac{1}{\tau_1 s+1}X(s), \tag{3.6}$$

а изображение выходной функции от начальных условий — функцией

$$Y_c(s)=\frac{\tau_1 y_0}{\tau_1 s+1}; \tag{3.7}$$

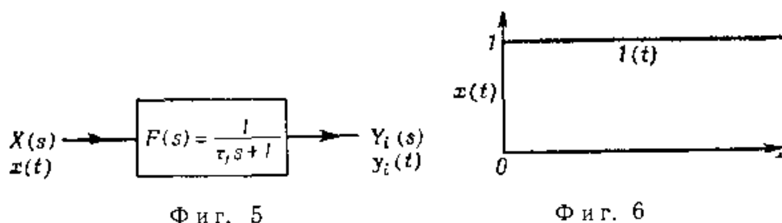
---

[1]) См. также [1], стр. 49, § 5 (вырожденная линейная система). — *Прим. перев.*

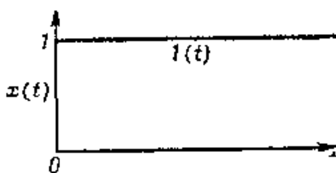
передаточная функция $F(s)$ имеет вид

$$F(s)=\frac{1}{\tau_1 s+1}\,. \tag{3.8}$$

Уравнению (3.6) можно дать графическое представление с помощью фиг. 5. Эта простая наглядная иллюстрация очень полезна для уяснения смысла передаточной функции и обычно называется блок-схемой.



Фиг. 5



Фиг. 6

Подсчитаем выход системы для нескольких различных случаев входной функции $x(t)$. Зададимся, прежде всего, входом $x(t)$ в виде единичной функции $1(t)$, изображенной графически на фиг. 6. Тогда

$$X(s)=\int_0^\infty e^{-st}dt=\frac{1}{s}$$

и

$$Y_i(s)=\frac{1}{s(\tau_1 s+1)}=\frac{1}{s}-\frac{1}{s+(1/\tau_1)}\,.$$

В соответствии со «словарем» оригиналов и изображений (табл. 1) выпишем выходную функцию от входного воздействия

$$y_i(t)=1-e^{-t/\tau_1}; \tag{3.9}$$

выходная функция от начальных условий, определяемая изображением (3.7), равна

$$y_c(t)=y_0 e^{-t/\tau_1}. \tag{3.10}$$

Графики этих выходных функций изображены на фиг. 7. Мы нашли, что выход от начальных условий представляет собой затухающую экспоненту с постоянной времени τ_1, а выход от входного воздействия — экспоненту,

также с постоянной времени τ_1, асимптотически приближающуюся к единичной функции. При $t=\tau_1$ выход $y_i(t)$ достигает 63% от его асимптотической предельной величины.

При этих условиях рассогласование $e(t)$ (или ошибка), определяемое как разность между входом $x(t)$ и выходом $y_i(t)$, равно

$$e(t)=x(t)-y_i(t)=e^{-t/\tau_1}. \qquad (3.11)$$

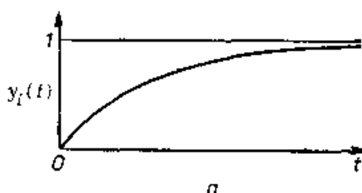
Поэтому рассогласование исчезает при $t \to \infty$.

Перейдем к другому примеру входного воздействия. Пусть входная величина изменяется по синусоидальному закону

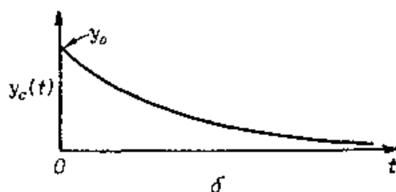
$$x(t)=x_m e^{i\omega t},$$

где x_m — амплитуда колебаний, а ω — их частота. Тогда

$$X(s)=\frac{x_m}{s-i\omega}. \qquad (3.12)$$



а



б

Фиг. 7

Выход от начальных условий имеет в данном случае тот же вид, что и в предыдущем примере. Выход же от входного воздействия имеет вид

$$Y_i(s) \doteq x_m \frac{1}{(s-i\omega)(\tau_1 s+1)}=\frac{x_m}{1+i\omega\tau_1}\left[-\frac{1}{s+(1/\tau_1)}+\frac{1}{s-i\omega}\right].$$

Следовательно, согласно «словарю» оригиналов и изображений,

$$y_i(t)=-\frac{x_m}{1+i\omega\tau_1}e^{-t/\tau_1}+\frac{x_m}{1+i\omega\tau_1}e^{i\omega t}.$$

Первое слагаемое правой части последнего равенства представляет собой чистую затухающую экспоненту, а второе слагаемое характеризует установившуюся составляющую полного решения системы. Таким образом,

$$\frac{[y(t)]_{\text{уст.}}}{x(t)}=\frac{1}{1+i\omega\tau_1}=F(i\omega).$$

Этот вывод находится в полном соответствии с нашим общим результатом, выраженным соотношением (2.16). Так как

$$\frac{1}{1+i\omega\tau_1} = \frac{1}{\sqrt{1+\omega^2\tau_1^2}}\, e^{-i\,\mathrm{arctg}\,\omega\tau_1}, \tag{3.13}$$

то установившуюся составляющую можно представить в виде

$$[y(t)]_{\text{уст.}} = \frac{x_m}{\sqrt{1+\omega^2\tau_1^2}}\, e^{i(\omega t - \mathrm{arctg}\,\omega\tau_1)}.$$

Следовательно, амплитуда установившейся составляющей выходной переменной системы уменьшается в $\sqrt{1+\omega^2\tau_1^2}$ раз по сравнению с амплитудой входной переменной, а фаза этой составляющей выходной переменной отстает от входной переменной на величину $\mathrm{arctg}\,\omega\tau_1$. При синусоидальных входах низкой частоты

$$[y(t)]_{\text{уст.}} \approx x_m e^{i\omega(t-\tau_1)}, \qquad \tau_1\omega \ll 1. \tag{3.14}$$

Таким образом, в этих условиях амплитуда не испытывает изменения по сравнению с амплитудой входа, но фаза установившейся составляющей выхода отстает от фазы входа на величину, равную постоянной времени τ_1 передаточной функции. Наоборот, при синусоидальных входах высокой частоты

$$[y(t)]_{\text{уст.}} \approx \frac{x_m}{\omega\tau_1}\, e^{i\left(\omega t - \frac{\pi}{2}\right)}, \quad \tau_1\omega \gg 1, \tag{3.15}$$

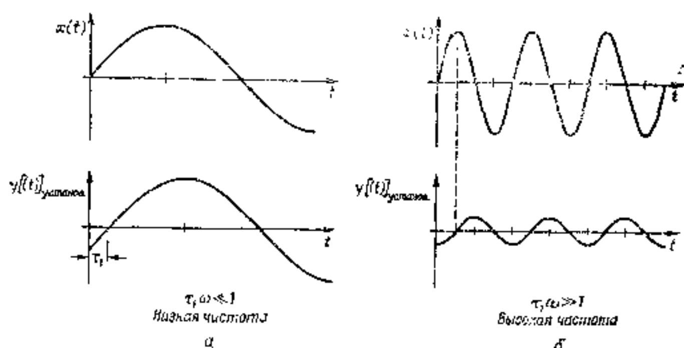
т. е. в этом случае амплитуда уменьшается в $\omega\tau_1$ раз, а запаздывание по фазе становится равным $\pi/2$. Графики входа и установившейся составляющей выхода в случаях низкочастотных и высокочастотных входных колебаний соответственно изображены на фиг. 8.

3.2. Способы задания передаточной функции. Передаточная функция $F(s)$ представляет собой функцию комплексной переменной s. Так как в общем случае функция $F(s)$ выражается отношением двух многочленов относительно s, то она определяется с точностью до постоянного множителя заданием ее нулей и полюсов. Для нахождения же постоянного множителя достаточно знать значение

$F(s)$ при одном произвольном значении s. Удобнее всего задавать значение $F(s)$ в начале координат. В самом деле, величина

$$|F(0)| = K \tag{3.16}$$

имеет определенный физический смысл: она равна отношению выхода к входу при постоянном входном воздействии, т. е. при входном сигнале, не меняющемся с течением времени. Отношение K называют коэффициентом



Фиг. 8

усиления системы. Поэтому передаточная функция определяется единственным образом с помощью нулей, полюсов и коэффициента усиления. Таков один из способов задания передаточной функции. Например, коэффициент усиления простой передаточной функции, определенной соотношением (3.8), равен единице; она имеет простой полюс в точке $-1/\tau_1$ и не имеет нулей.

Если заданы одновременно значения как вещественной, так и мнимой частей функции $F(s)$ вдоль мнимой оси плоскости комплексной переменной s, то на основании принципа аналитического продолжения [1]) тем самым функция $F(s)$ определена и при любом s. Поэтому другой возможный способ определения $F(s)$ заключается в задании функции $F(i\omega)$ комплексной переменной $i\omega$ (ω — веществен-

[1]) См. [14, 17—19]. — *Прим. перев.*

ная переменная). В физических системах все коэффициенты при степенях s многочлена $N(s)$, представляющего собой числитель передаточной функции $F(s)$, и все коэффициенты при степенях s многочлена $D(s)$, представляющего собой знаменатель $F(s)$, суть вещественные числа. Поэтому, если обозначить через $\overline{F}$ комплексное число, сопряженное с комплексным числом F, то

$$F(-i\omega) = \overline{F}(\overline{i\omega}). \tag{3.17}$$

Таким образом, в физических системах для определения функции $F(s)$ комплексной переменной s при любых значениях s достаточно знать функцию $F(i\omega)$ при $\omega \geqslant 0$. Но, в силу соотношения (2.16), $F(i\omega)$ представляет собой отношение установившейся составляющей выходной переменной к синусоидальной входной переменной частоты ω. Функция $F(i\omega)$, заданная при всех значениях ω, называется частотной характеристикой системы [1]). Таким образом, с помощью частотной характеристики осуществляется еще один способ задания передаточной функции. В разобранной выше простой системе первого порядка частотная характеристика определена функцией (3.13).

Один из способов представления частотной характеристики, разработанный Г. У. Боде (H. W. Bode), называется диаграммой Боде. Обозначим через M модуль функции $F(i\omega)$, а через θ – ее аргумент, т. е.

$$F(i\omega) = Me^{i\theta}. \tag{3.18}$$

Диаграмма Боде заключается в построении графиков $\lg M$ и θ в функции от $\lg \omega$ [2]). Причина, по которой модуль M откладывается в логарифмической шкале, в то время как фазовый угол θ откладывается в натуральную величину, будет выяснена в дальнейшем. Для нашей простой си-

---

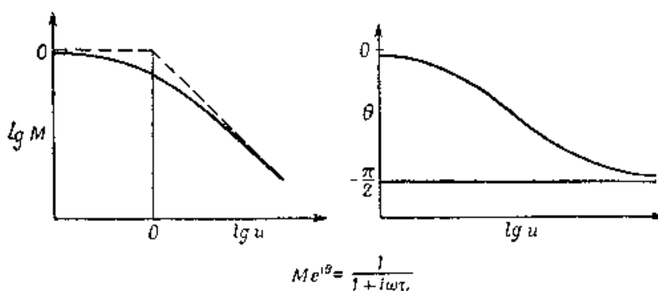
[1]) Иногда функцию $F(i\omega)$ называют передаточной функцией, не делая различия в названии между $F(s)$ и $F(i\omega)$, или же еще амплитудной характеристикой. — *Прим. перев.*

[2]) При определении диаграммы Боде автор не указывает основания логарифмов, ибо оно не имеет принципиального значения. В приложениях же всюду применяются десятичные логарифмы (см. фиг. 9 и т. д.). — *Прим. перев.*

стемы, определяемой частотной характеристикой (3.13),

$$\left.\begin{aligned} M &= \frac{1}{\sqrt{1+\omega^2\tau_1^2}} = \frac{1}{\sqrt{1+u^2}}, \\ \theta &= -\operatorname{arctg}\omega\tau_1 = -\operatorname{arctg} u, \end{aligned}\right\} \quad u = \omega\tau_1, \qquad (3.19)$$

где u – безразмерная частота [1]). Диаграмма Боде для этой системы изображена на фиг. 9. Выводы относительно поведения частотной характеристики в областях низких и высоких частот, вытекающие из рассмотрения диаграммы



$Me^{i\theta} = \frac{1}{1+i\omega\tau_1}$

Фиг. 9

Боде, совпадают с заключениями, полученными ранее в виде соотношений (3.14) и (3.15). При $u \to \infty$ угловой коэффициент графика $\lg M$ в функции $\lg u$ стремится к -1. При малых значениях u этот коэффициент близок к нулю. Поэтому график $\lg M$ по $\lg u$ для системы первого порядка можно аппроксимировать двумя прямыми линиями [2]).

В литературе по акустике и электротехнике установилась практика откладывать по оси ординат $20 \lg M$ вместо $\lg M$, что делается с целью выражения единиц измерения амплитуды в децибелах. Интервал частот, заключенный между данным значением частоты и удвоенным значением этой частоты, называется октавой, следо-

---

[1]) Зависимость $\lg M$ от $\lg \omega$ (или $\lg u$) обычно называется логарифмической амплитудной характеристикой системы, а зависимость θ от $\lg \omega$ — ее логарифмической фазовой характеристикой [9,21]. — *Прим. перев.*

[2]) Ломаная на фиг. 9, нанесенная пунктиром, называется асимптотической или (трапецоидальной) логарифмической амплитудной характеристикой [9]. — *Прим. перев.*

вательно, та часть графика, в которой угловой коэффициент аппроксимированной диаграммы равен -1, характеризуется наклоном этой диаграммы в $-20 \lg 2 = -6{,}02$ *дб* на октаву. Отметим также, что на этой же фигуре аппроксимированная диаграмма (пунктирная линия) проходит через 0 при $u = 1$, т. е. при $\omega = 1/\tau_1$. Поэтому мы можем снять частотную характеристику системы первого порядка экспериментально и нанести указанным способом результаты измерения в виде диаграммы. Постоянную τ_1 времени системы нетрудно оценить с помощью значения частоты, при которой наклонная часть аппроксимированной диаграммы, соответствующая высоким частотам ω, пересекает горизонтальную ось.

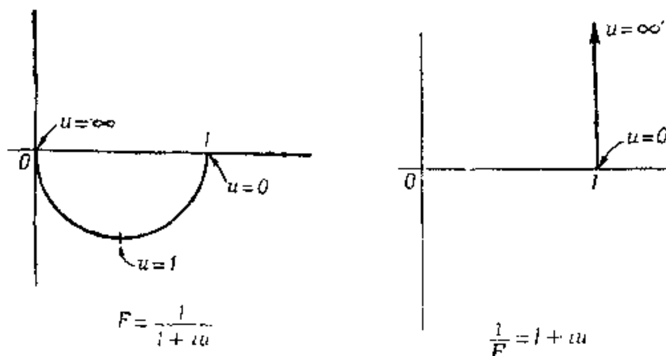
Другой способ представления частотной характеристики был разработан Г. Найквистом и получил название диаграммы Найквиста. Этот способ заключается в непосредственном построении графиков комплексной функции $F(i\omega)$ [или $1/F(i\omega)$] на плоскости комплексной переменной F (или $1/F$ соответственно). При этом параметром кривой служит частота ω. Для простой системы, рассмотренной выше, т. е. системы первого порядка, график функции $F(i\omega)$ представляет собой полуокружность, начинающуюся в точке 1 при $\omega = 0$, проходящую через точку $1/(1+i) = (1/\sqrt{2})(1-i)$ при $\omega\tau_1 = u = 1$ и оканчивающуюся в начале координат при $\omega \to \infty$. График функции $1/F$ намного проще: $1/F = 1 + i\omega\tau_1$, и поэтому этот график сводится к прямой линии, параллельной мнимой оси. Обе диаграммы Найквиста для системы первого порядка изображены на фиг. 10[1]).

3.3. Примеры систем первого порядка. Сложные системы содержат много элементов, которые можно приближенно описать с помощью передаточной функции первого порядка. Мы рассмотрим кратко несколько примеров таких элементов вместе с диаграммами их частотных характеристик.

---

[1]) Функция $1/F$ называется иногда обратной амплитудно-фазовой характеристикой. График $F(i\omega)$ на плоскости F называется, как правило, амплитудно-фазовой характеристикой. —*Прим. перев.*

Интегрирующее звено. Электрический двигатель, в котором угловая скорость $d\varphi/dt$ пропорциональна входному



Фиг. 10

напряжению v, описывается уравнением

$$\frac{d\varphi}{dt} = Kv, \tag{3.20}$$

где K — коэффициент пропорциональности[1]). Таким образом, угловая координата φ двигателя пропорциональна интегралу

$$\int_0^t v\,dt.$$

Действие электрического двигателя можно охарактеризовать с помощью блок-схемы, показанной на фиг. 11, где через $V(s)$ и $\Phi(s)$ обозначены изображения переменных v и φ соответственно.

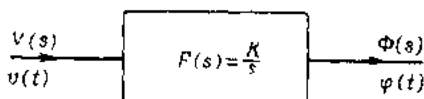
Передаточная функция двигателя $F(s) = K/s$ представляет собой предельное значение функции $1/(\tau_1 s + 1)$ при $\tau \to \infty$ и характеризуется простым полюсом в начале координат.

Чтобы рассматривать постоянную K в качестве коэффициента усиления системы в выражении передаточной

---

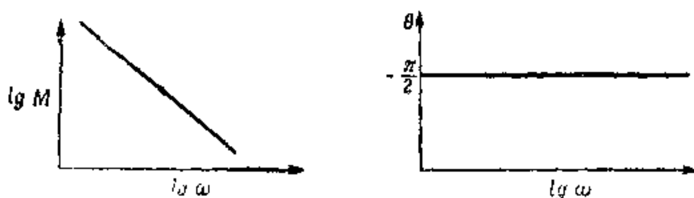
[1]) Уравнение (3.20) приближенно справедливо, например, для двигателя постоянного тока независимого возбуждения, находящегося в режиме холостого хода. — *Прим. ред.*

функции $F(s)$, мы должны видоизменить определение коэффициента усиления, введенное в предыдущем разделе. Данное там определение пригодно для передаточных функций, не имеющих в начале ни нулей, ни полюсов. Поэтому



Фиг. 11

коэффициент усиления K интегрирующего звена, т. е. звена, передаточная функция которого имеет при $s=0$



$F(i\omega)=\frac{K}{i\omega}=Me^{i\theta}$

Фиг. 12

простой полюс, следует определить с помощью предельного перехода

$$K=\lim_{s\to 0}|sF(s)|. \tag{3.21}$$

Выпишем частотную характеристику интегрирующего звена

$$F(i\omega)=\frac{K}{i\omega}=\left(\frac{K}{\omega}\right)e^{-i\frac{\pi}{2}}.$$

Поэтому, в силу (3.18),

$$M=\frac{K}{\omega}, \qquad \theta=-\frac{\pi}{2}. \tag{3.22}$$

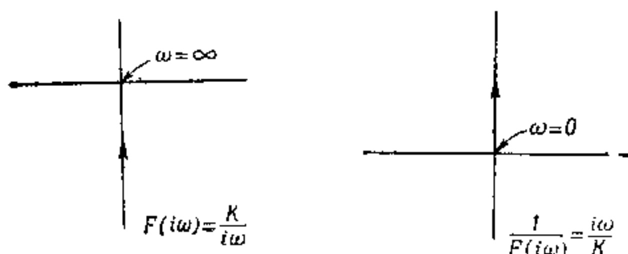
Соответствующая диаграмма Боде показана на фиг. 12, а диаграмма Найквиста — на фиг. 13.

Дифференцирующее звено. Скоростной гироскоп выдает выходное напряжение v, пропорциональное угловой ско-

рости $d\varphi/dt$ прецессии, т. е.

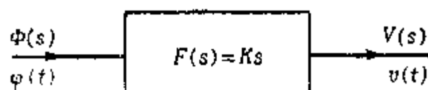
$$v = K\frac{d\varphi}{dt},$$

где K — коэффициент пропорциональности. Действие этого звена противоположно действию только что рассмотренного звена. Передаточная функция $F(s) = Ks$ имеет нуль



Фиг. 13

в начале координат. Таким образом, коэффициент усиления дифференцирующего звена, т. е. системы, передаточ-



Фиг. 14

ная функция $F(s)$ которой имеет нуль при $s = 0$, следует определить с помощью предельного перехода

$$K = \lim_{s\to 0}\left|\frac{F(s)}{s}\right|. \tag{3.23}$$

Блок-схема дифференцирующего звена изображена на фиг. 14, его диаграмма Боде — на фиг. 15, а диаграмма Найквиста — на фиг. 16.

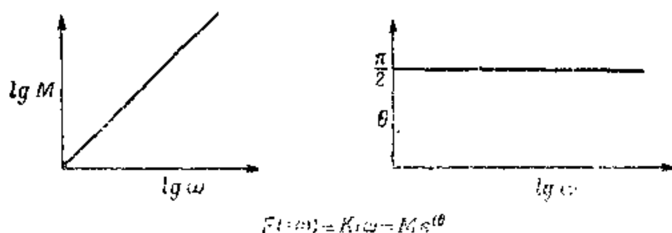
Простое фазосдвигающее звено. Рассмотрим электрическую цепь, состоящую из сопротивления и емкости (фиг. 17). Пусть j — ток, текущий через сопротивление R, и C — емкость. Если в начальный момент $(t = 0)$

конденсатор разряжен, то

$$jR + \frac{1}{C}\int_0^t j(t)\,dt = v_1,$$

$$\frac{1}{C}\int_0^t j(t)\,dt = v_2.$$

Умножив обе части каждого из этих равенств на e^{-st}

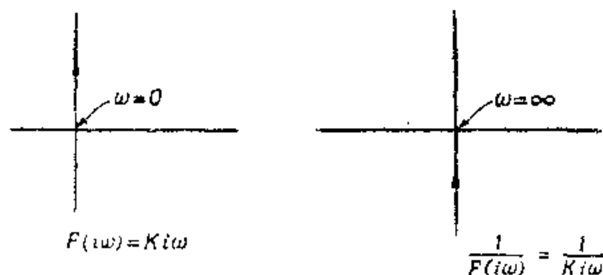


Фиг. 15

и проинтегрировав от $t=0$ до $t=\infty$, получим

$$\left(R + \frac{1}{Cs}\right) J(s) = V_1(s),$$

$$\frac{1}{Cs} J(s) = V_2(s).$$

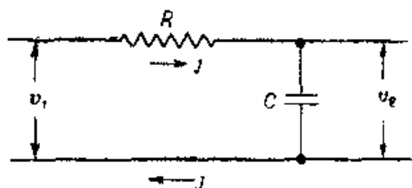


Фиг. 16

Поэтому

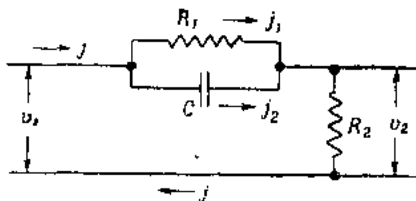
$$\frac{V_2(s)}{V_1(s)} = F(s) = \frac{1}{1+RCs}. \tag{3.24}$$

Отсюда следует, что передаточная функция этой цепи типа RC имеет такой же вид, как и передаточная функция системы со скользящей пружиной и катарактом; ее



Фиг. 17

постоянная времени $\tau_1 = RC$. Поэтому диаграммы Боде и Найквиста для этой системы изображаются диаграммами на фиг. 9 и 10 соответственно. Фазосдвигающее



Фиг. 18

звено часто применяется для введения в систему отставания по фазе[1]).

Фазоопережающее звено. На фиг. 18 показана более сложная электрическая цепь. Ее уравнения имеют вид

$$j = j_1 + j_2,$$

$$R_1 j_1 = \frac{1}{C} \int_0^t j_2(t)\, dt$$

$$v_1 = R_1 j_1 + R_2 j,$$

$$v_2 = R_2 j.$$

[1]) Это звено иногда называется инерционным или релаксационным звеном.— *Прим. ред.*

Соответствующие уравнения в изображениях будут

$$J = J_1 + J_2,$$

$$R_1 J_1 = \frac{1}{Cs} J_2$$

и

$$V_1 = R_1 J_1 + R_2 J,$$

$$V_2 = R_2 J.$$

Поэтому

$$\frac{V_2(s)}{V_1(s)} = F(s) = \frac{R_2 + R_1 R_2 Cs}{(R_1 + R_2) + R_1 R_2 Cs}.$$

Отсюда находим K:

$$K = \frac{R_2}{R_1 + R_2} = r. \tag{3.25}$$

Коэффициент усиления непременно выражается числом, меньшим единицы; обычно оно заключено между 0,1 и 1. Введя постоянную ω_1 с помощью соотношения

$$\omega_1 = \frac{R_1 + R_2}{R_1 R_2 C}, \tag{3.26}$$

представим передаточную функцию в виде

$$F(s) = r\,\frac{1 + (s/r\omega_1)}{1 + (s/\omega_1)}. \tag{3.27}$$

Эта функция имеет нуль в точке $-r\omega_1$ и полюс в точке $-\omega_1$[1]).

Частотная характеристика фазоопережающего звена выражается функцией

$$F(i\omega) = \frac{r\omega_1 + i\omega}{\omega_1 + i\omega}. \tag{3.28}$$

Введя безразмерную частоту u с помощью соотношения

$$u = \frac{1}{\sqrt{r}} \cdot \frac{\omega}{\omega_1}, \tag{3.29}$$

---

[1]) Это звено иногда называют форсирующим звеном. — *Прим. ред.*

получим

$$M=\sqrt{r}\sqrt{\frac{1+\frac{u^2}{r}}{\frac{1}{r}+u^2}}; \qquad \theta=\operatorname{arctg}\frac{u}{\sqrt{r}}-\operatorname{arctg}(\sqrt{r}u). \quad (3.30)$$

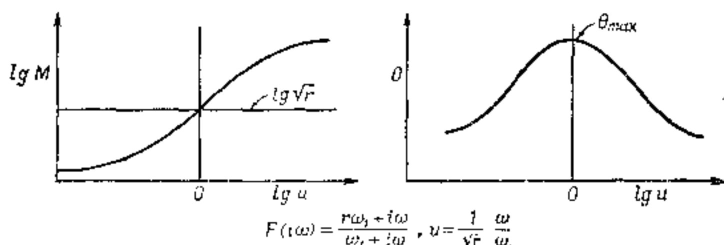
Отсюда

$$\lg M(u)=\lg\sqrt{r}+\lg\sqrt{\frac{1+\frac{u^2}{r}}{\frac{1}{r}+u^2}}=\lg\sqrt{r}-\lg\sqrt{\frac{1+\frac{1}{ru^2}}{\frac{1}{r}+\frac{1}{u^2}}}$$

и

$$\theta(u)=\theta\left(\frac{1}{u}\right).$$

Диаграмма Боде для этого звена отличается некоторой



$F(i\omega)=\frac{r\omega_1+i\omega}{\omega_1+i\omega}$, $u=\frac{1}{\sqrt{r}}\frac{\omega}{\omega_1}$

Фиг. 19

симметрией относительно прямой $u=1$ (фиг. 19). Максимум фазы θ имеет место при $u=1$ и равен

$$\theta_{max}=\operatorname{arctg}\frac{1}{\sqrt{r}}-\operatorname{arctg}\sqrt{r}=\frac{\pi}{2}-2\operatorname{arctg}\sqrt{r}. \quad (3.31)$$

Поэтому в определенной области частот данная цепь доставляет значительное опережение по фазе. При очень больших ω имеет место равенство $M=1$, при очень малых ω — равенство $M=r$.

Полосовое фазосдвигающее звено. Передаточная функция RC электрической цепи, изображенной на фиг. 20, имеет вид

$$\frac{V_2(s)}{V_1(s)}=F(s)=\frac{1+R_1Cs}{1+(R_1+R)Cs}.$$

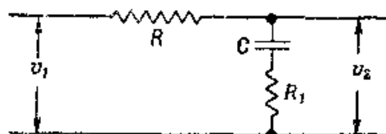
Усиление этой цепи равно единице. Введя постоянные

$$\omega_1 = \frac{1}{R_1 C}, \qquad r = \frac{R_1}{R + R_1}, \tag{3.32}$$

перепишем передаточную функцию в виде

$$F(s) = \frac{1 + \frac{s}{\omega_1}}{1 + \frac{s}{r\omega_1}}. \tag{3.33}$$

Сравнив последнее выражение с выражением (3.28) частотной характеристики фазоопережающего звена, мы убедимся, что передаточные функции этих звеньев суть взаимно



Фиг. 20

обратные функции. В самом деле, частотную характеристику рассматриваемого звена можно представить с помощью отношения

$$F(i\omega) = \frac{1 + i\frac{\omega}{\omega_1}}{1 + i\frac{\omega}{r\omega_1}} = \frac{1 + i\sqrt{r}\,u}{1 + i\frac{1}{\sqrt{r}}u},$$

где безразмерная частота u определена равенством

$$u = \frac{1}{\sqrt{r}}\frac{\omega}{\omega_1}. \tag{3.34}$$

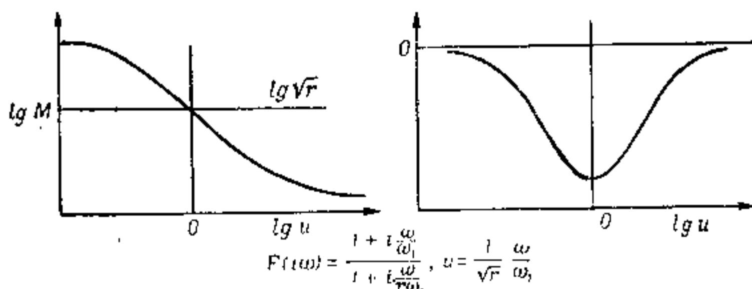
Таким образом,

$$M = \sqrt{r}\sqrt{\frac{\frac{1}{r} + u^2}{1 + \frac{u^2}{r}}}; \quad \theta = \operatorname{arctg}(\sqrt{r}\,u) - \operatorname{arctg}\frac{u}{\sqrt{r}}. \tag{3.35}$$

Соответствующая диаграмма Боде изображена на фиг. 21. В некоторой полосе частот наблюдается запаздывание по фазе. Наибольшее запаздывание по фазе достигается при

$u=1$, т. е. при $\omega=\sqrt{r}\,\omega_1$, а его величина определяется соотношением (3.31).

Упрощенное описание движения самолета по крену. Введем обозначения: I — момент инерции самолета относительно его продольной оси, φ — угол крена, L_p — коэффициент аэродинамического демпфирования в движении



$$F(i\omega)=\frac{1+i\frac{\omega}{\omega_1}}{1+i\frac{\omega}{r\omega_1}},\quad u=\frac{1}{\sqrt{r}}\,\frac{\omega}{\omega_1}$$

Фиг. 21

самолета по крену, $k\delta$ — момент, приложенный к самолету со стороны элеронов, отклоненных на угол δ. Тогда крен самолета описывается уравнением

$$I\frac{d^2\varphi}{dt^2}+L_p\frac{d\varphi}{dt}=k\delta.$$

Обозначим через $p=d\varphi/dt$ угловую скорость крена; в соответствии с этим обозначением предыдущее уравнение примет вид

$$I\frac{dp}{dt}+L_p p=k\delta.$$

Пусть в момент времени $t=0$ угловая скорость крена равна нулю. Применив к уравнению движения самолета по крену преобразование Лапласа, получим уравнение в изображениях

$$(Is+L_p)\,P(s)=k\Delta(s).$$

Таким образом, для передаточной функции самолета при его колебаниях по крену мы получаем выражение

$$\frac{P(s)}{\Delta(s)}=F(s)=\frac{k}{Is+L_p}=\frac{k}{L_p}\,\frac{1}{1+\frac{I}{L_p}s} \tag{3.36}$$

Следовательно, поведение кренящегося самолета в той мере, в какой оно описывается этой передаточной функцией, подобно поведению скользящей пружины с катарактом и простого фазосдвигающего контура[1]). Постоянная времени τ_1 здесь равна $\frac{I}{L_p}$. При очень слабом демпфировании $\tau_1 \to \infty$ и система ведет себя как простой интегратор.

3.4. Системы второго порядка. Обратимся опять к системе скользящей пружины с катарактом (фиг. 4), но введем массу m на конце пружины, примыкающем к катаракту. Добавление массы приведет к появлению силы инерции $m\, d^2y/dt^2$, и уравнение движения примет вид

$$m\frac{d^2y}{dt^2}+c\frac{dy}{dt}+ky=kx$$

при начальных условиях

$$\left.\begin{aligned} y(0)&=y_0, \\ \left(\frac{dy}{dt}\right)_{t=0}&=y_0^{(1)}. \end{aligned}\right\} \tag{3.37}$$

Дифференциальное уравнение движения можно переписать в более удобном виде с помощью обозначений

$$\omega_0^2=\frac{k}{m}\,; \qquad \zeta=\frac{\left(\frac{c}{m}\right)}{2\omega_0}. \tag{3.38}$$

Здесь ω_0 — собственная частота системы масса — пружина при отсутствии катаракта, а ζ — безразмерный коэффициент демпфирования, т. е. отношение истинного коэффициента демпфирования к его критическому значению; смысл этого параметра выявится несколько ниже. В новых обозначе-

---

[1]) То же заключение справедливо в отношении рысканья (относительно заданного курса) горизонтально летящего самолета. Колебания самолета при тангаже уже не описываются передаточной функцией (3.36), ибо в этом случае на самолет действует аэродинамический момент, зависящий от угла атаки, тогда как в движениях крена и рысканья естественная направляющая сила отсутствует. — *Прим. перев.*

ниях дифференциальное уравнение системы приводится к виду

$$\frac{d^2y}{dt^2} + 2\zeta\omega_0 \frac{dy}{dt} + \omega_0^2 y = \omega_0^2 x. \qquad (3.39)$$

Применив к уравнению (3.39) преобразование Лапласа с учетом начальных условий (3.37), мы получим следующее уравнение в области изображений:

$$(s^2 + 2\zeta\omega_0 s + \omega_0^2)\, Y(s) = \omega_0^2 X(s) + y_0^{(1)} + (s + 2\zeta\omega_0)\, y_0.$$

Таким образом, выход системы от начальных условий будет иметь вид

$$Y_c(s) = \frac{y_0 s + (y_0^{(1)} + 2\zeta\omega_0 y_0)}{s^2 + 2\zeta\omega_0 s + \omega_0^2}, \qquad (3.40)$$

а для передаточной функции системы мы получим выражение

$$F(s) = \frac{Y_i(s)}{X(s)} = \frac{1}{\left(\frac{s}{\omega_0}\right)^2 + 2\zeta\frac{s}{\omega_0} + 1}. \qquad (3.41)$$

Итак, система обладает коэффициентом усиления $K = 1$, а ее передаточная функция не имеет нулей и имеет два простых полюса в точках

$$\left.\begin{aligned} \frac{s_1}{\omega_0} &= -\zeta + \sqrt{\zeta^2 - 1}, \\ \frac{s_2}{\omega_0} &= -\zeta - \sqrt{\zeta^2 - 1}, \end{aligned}\right\} \quad \zeta^2 > 1. \qquad (3.42)$$

Если коэффициент демпфирования катаракта меньше критического, то величина ζ меньше единицы. В этом случае полюсы s_1 и s_2 комплексные сопряженные величины. Обозначив через λ и ν соответственно вещественную и мнимую части полюса s_1, напишем

$$\left.\begin{aligned} \frac{s_1}{\omega_0} &= -\zeta + i\sqrt{1 - \zeta^2} = \frac{\lambda + i\nu}{\omega_0} = e^{i\varphi_1}, \\ \frac{s_2}{\omega_0} &= -\zeta - i\sqrt{1 - \zeta^2} = \frac{\lambda - i\nu}{\omega_0} = e^{-i\varphi_1}, \end{aligned}\right\} \quad \zeta^2 < 1, \qquad (3.43)$$

где последняя форма представления полюсов возможна благодаря равенству модулей обоих этих полюсов. При положительном демпфировании λ представляет собой отрицательное число.

С помощью соотношения (3.40) нетрудно определить выход $y_c(t)$ системы от начальных условий. Так, при $\zeta^2 < 1$ полюсы передаточной функции выражаются равенствами (3.43), и мы найдем, что

$$y_c(t) = \frac{y_0^{(1)}}{\nu} e^{\lambda t} \sin \nu t + y_0 e^{\lambda t} \cos \nu t + \frac{-\lambda}{\nu} y_0 e^{\lambda t} \sin \nu t. \quad (3.44)$$

Так как число λ отрицательное, то выход системы представляет собой затухающее движение, выражающееся синусоидальной функцией. При $\zeta^2 > 1$ выход системы сводится к апериодически затухающему движению. Таким образом, при демпфировании, превышающем критическое, выход $y_c(t)$ имеет неколебательный характер. В этом состоит смысл понятия критического демпфирования.

Зададим теперь вход $x(t)$ системы в форме единичной функции $1(t)$, изображенной на фиг. 6. Тогда $X(s) = 1/s$. При $\zeta^2 < 1$

$$Y_i(s) = \frac{\omega_0^2}{s[(s-\lambda)^2 + \nu^2]}$$

и, следовательно, выход $y_i(t)$ от входного воздействия выражается функцией

$$y_i(t) = 1 - \left[\cos \nu t + \left(\frac{-\lambda}{\nu}\right) \sin \nu t\right] e^{\lambda t}. \quad (3.45)$$

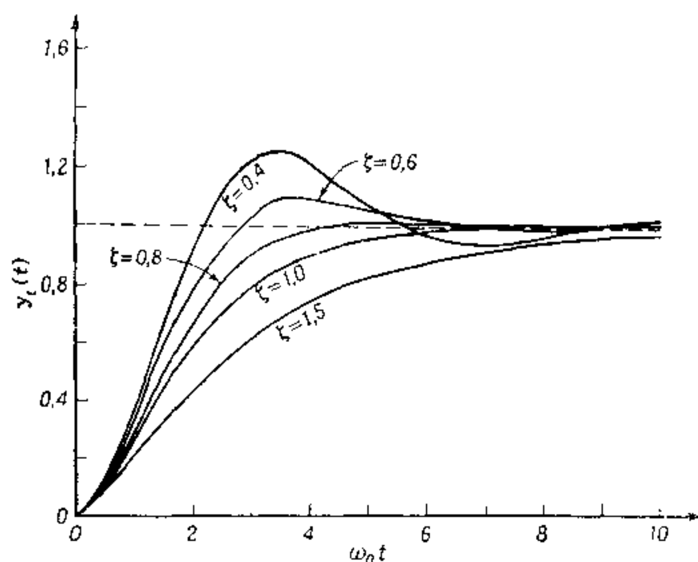
При $\zeta^2 > 1$ этот выход представляет собой неколебательную функцию, равную

$$y_i(t) = 1 - \frac{1}{2\sqrt{\zeta^2 - 1}} \left[\frac{e^{s_1 t}}{\zeta - \sqrt{\zeta^2 - 1}} - \frac{e^{s_2 t}}{\zeta + \sqrt{\zeta^2 - 1}}\right], \quad (3.46)$$

где s_1 и s_2 определены выражениями (3.42). Течение выходной функции $y_i(t)$, определяемое последним соотношением, графически изображается семейством кривых на фиг. 22, соответствующих различным значениям безразмерного коэффициента демпфирования ζ. Рассматривая течение этих кривых, видим, что для обеспечения быстрого приближения выходной функции системы к ее асимптотическому значению следует выбирать не слишком большую величину ζ.

С другой стороны, если ζ слишком мало, то в системе будут наблюдаться колебания со значительным перерегулированием. Поэтому необходимо принять некоторое

компромиссное решение. Обычно на практике значение безразмерного коэффициента демпфирования ζ выбирают между 0,4 и 1.



Фиг. 22

Если на вход системы подается синусоидальное колебание (3.11) амплитуды x_m и частоты ω, то

$$Y_i(s) = \frac{x_m}{s - i\omega} F(s) = \frac{x_m}{s - i\omega} \cdot \frac{\omega_0^2}{s^2 + 2\zeta\omega_0 s + \omega_0^2}.$$

Поэтому при $\zeta^2 < 1$ выход системы от входного воздействия имеет вид

$$y_i(t) = x_m F(i\omega) e^{i\omega t} + \frac{x_m}{2i\nu} \cdot \frac{\omega_0^2}{\lambda + i(\nu - \omega)} e^{(\lambda + i\nu) t} -$$

$$- \frac{x_m}{2i\nu} \cdot \frac{\omega_0^2}{\lambda - i(\nu + \omega)} e^{(\lambda - i\nu) t}, \qquad (3.47)$$

где λ и ν определены выражениями (3.43). Так как при положительном демпфировании величина λ отрицательна, то в установившемся режиме выход системы характери-

зуется первым слагаемым правой части равенства (3.47), что отвечает общему выражению (2.16).

С помощью равенства (3.41) найдем частотную характеристику рассматриваемой системы второго порядка в виде

$$F(i\omega) = Me^{i\theta} = \frac{1}{\left[1-\left(\frac{\omega}{\omega_0}\right)^2\right] + 2i\zeta\left(\frac{\omega}{\omega_0}\right)}.$$

Поэтому[1])

$$M = \frac{1}{\sqrt{\left[1-\left(\frac{\omega}{\omega_0}\right)^2\right]^2 + \left[2\zeta\left(\frac{\omega}{\omega_0}\right)\right]^2}} \tag{3.48}$$

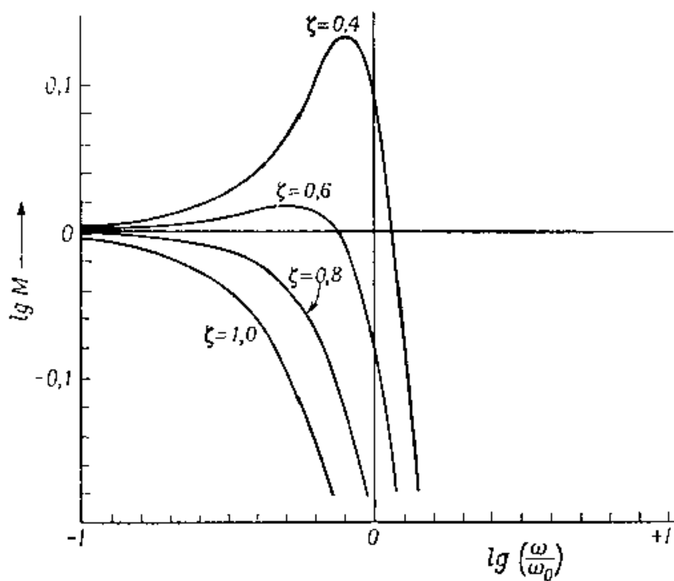
$$\operatorname{tg}\theta = -\frac{2\zeta\left(\frac{\omega}{\omega_0}\right)}{1-\left(\frac{\omega}{\omega_0}\right)^2}.$$

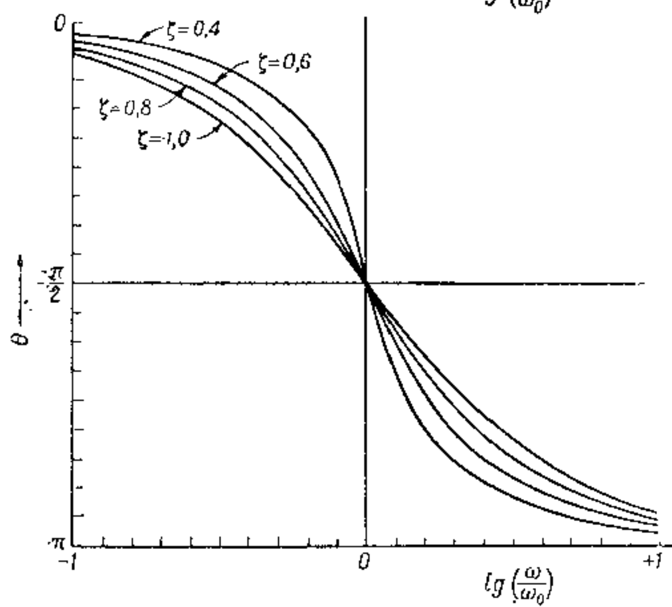
Соответствующая диаграмма Боде построена на фиг. 23. Максимальное значение M имеет место вблизи $\omega/\omega_0 = 1$, где $M \approx \frac{1}{2}\zeta$ и $\theta \approx -\pi/2$. При $\omega/\omega_0 \to \infty$ фазовый угол $\theta \to -\pi$ и $M \sim \frac{1}{(\omega/\omega_0)^2}$, т. е. $\ln M \sim -2\ln(\omega/\omega_0)$. Специалист в области акустики сказал бы, что в области высоких частот наклон этой характеристики составляет $-12{,}04$ *дб* на октаву.

Кривая Найквиста для этой системы построена на фиг. 24.

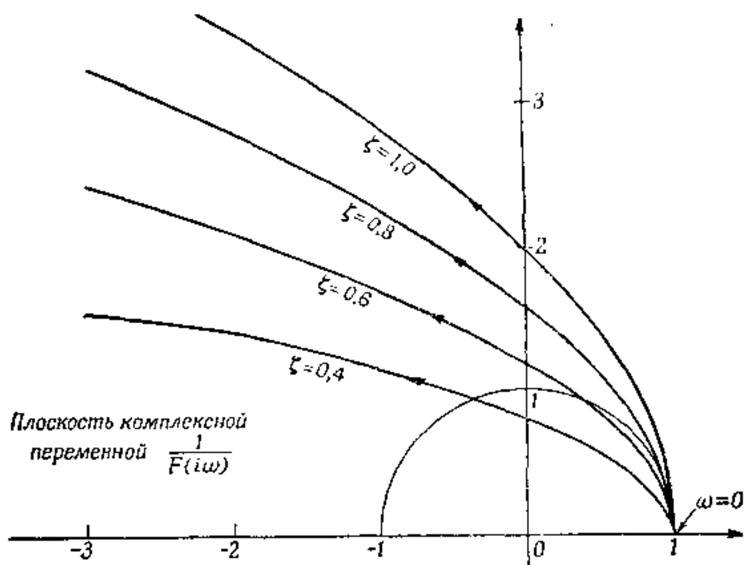
Посредством передаточной функции второй степени можно аппроксимировать и другие физические системы. Одним из примеров таких систем служит гидравлический серводвигатель. С помощью передаточной функции такого типа достигается и лучшее приближение для динамики скоростного гироскопа, разобранного в п. 3.3, а именно поведение этого гироскопа описывается посредством пере-

[1]) Метод оценки качества следящей системы с помощью чисел М и θ (амплитудно-фазовые диаграммы) рассмотрен, например, в [9, 21]. — *Прим. перев.*

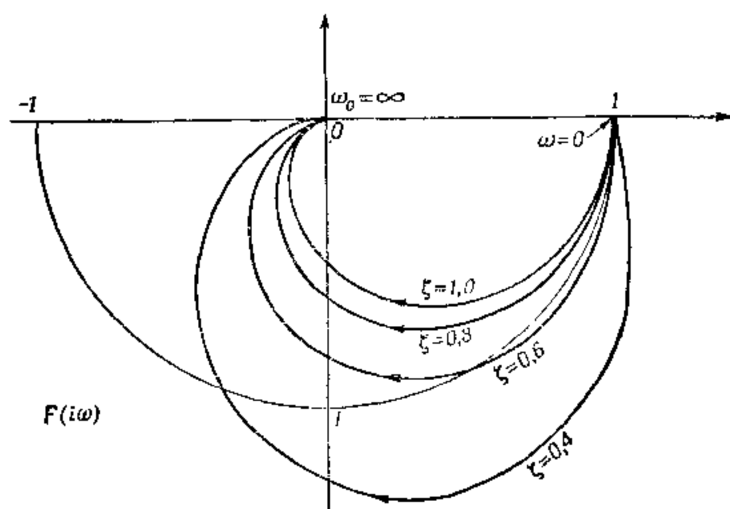

ζ=0,4
0,1
ζ=0,6
0
lg M
ζ=0,8
ζ=1,0
-0,1
-1
0
+1
lg $\left(\frac{\omega}{\omega_0}\right)$


0
ζ=0,4
ζ=0,6
ζ=0,8
ζ=1,0
θ
$-\frac{\pi}{2}$
-π
-1
0
+1
lg $\left(\frac{\omega}{\omega_0}\right)$

Фиг. 23



ζ=1,0
ζ=0,8
ζ=0,6
ζ=0,4
Плоскость комплексной переменной $\frac{1}{F(i\omega)}$
ω=0
-3
-2
-1
0
1
2
3



-1
$\omega_0=\infty$
0
1
ω=0
ζ=1,0
ζ=0,8
ζ=0,6
ζ=0,4
F(iω)

Фиг. 24

даточной функции

$$F(s) = \frac{Ks}{\left(\frac{s}{\omega_0}\right)^2 + 2\zeta\left(\frac{s}{\omega_0}\right) + 1}.$$

Представляет интерес сравнить эту более точную передаточную функцию с ее грубым приближением, показанным на фиг. 14. Отметим также передаточную функцию акселерометра

$$F(s) = \frac{Ks^2}{\left(\frac{s}{\omega_0}\right)^2 + 2\zeta\left(\frac{s}{\omega_0}\right) + 1}.$$

Таким же образом можно уточнить передаточную функцию электрического двигателя, используемого в качестве интегратора[1]):

$$F(s) = \frac{K}{s(\tau_1 s + 1)}.$$

Более грубое приближение этой функции было приведено на фиг. 11. Все рассмотренные здесь передаточные функции имеют то общее, что их знаменатели суть квадратные трехчлены. Смысл постоянных, фигурирующих в этих передаточных функциях, определяется аналогично тому, как это делалось в предшествующих примерах.

3.5. Определение частотной характеристики. В предыдущих пунктах мы рассматривали задачу об аналитическом нахождении передаточной функции $F(s)$ системы и ее частотной характеристики $F(i\omega)$, исходя из основных физических законов и зная структуру системы. Таким образом, эта методика подсчета частотной характеристики является аналитической, и достигаемая с ее помощью степень точности зависит от достоверности наших сведений об этой системе. Но в инженерной практике

---

[1]) То есть в том случае, когда выходной величиной считается не угловая скорость вала ротора, а угловое перемещение этого или жестко связанного с ним вала. — *Прим. ред.*

наши сведения, касающиеся подробностей строения рассматриваемой системы, часто оказываются неполными; иногда же, хотя строение системы нам и известно исчерпывающим образом, эта система оказывается настолько сложной, что аналитическое определение ее частотной характеристики перестает быть практически доступным по причине слишком громоздких вычислений. В таких случаях часто возникает необходимость определения частотной характеристики опытным путем. Наиболее простой метод экспериментального снятия частотной характеристики основан на том результате, что отношение установившегося выхода, возбуждаемого синусоидальным входом частоты ω, к этому входу и представляет собой частотную характеристику системы, как это следует из соотношения (2.16). Отношение амплитуды выхода к амплитуде входа равно числу M, а разность между фазой выхода и фазой входа — числу θ[1]). Поэтому этот экспериментальный метод требует определения отношений амплитуд и разностей фаз для некоторого числа частот в интересующем нас диапазоне частот. Этот метод снятия частотных характеристик применялся к различным задачам, начиная от столь простой системы, как топливный насос[2]), и вплоть до достаточно сложной системы, характеризующей продольное движение самолета[3]) [4]). Недостаток такого метода состоит в необходимости проведения длительных экспериментальных исследований в общем случае в широкой полосе частот. Кроме того, иногда приходится встречаться с затруднениями в измерении разности фаз между выходом и входом.

Более эффективный метод исследования состоит в том, чтобы возбуждать систему одновременно по всем частотам, а не по одной определенной частоте каждый раз отдельно. Лучше всего это сделать путем подачи на вход исследуемой системы единичного импульса. Тогда, согласно (2.18), переходный процесс в устойчивых системах

---

[1]) См. примечание на стр. 50. — *Прим. перев.*

[2]) Shames H., Himmel C. S., Blivas D., NACA TN 2109 (1950).

[3]) Milliken W. F., Journ. Aeronaut. Sci., 14, 493 (1947).

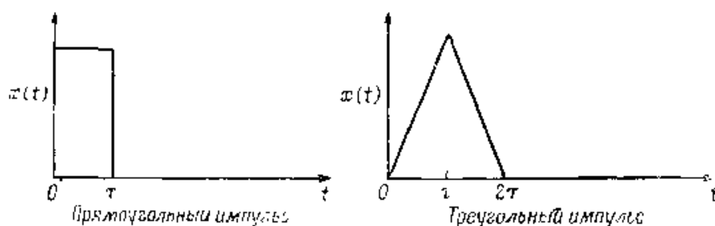
[4]) См. также [22—26]. — *Прим. перев.*

($\gamma=0$) определяется выражением

$$h(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}F(i\omega)\,e^{i\omega t}\,d\omega=$$

$$=\frac{1}{\pi}\int_{0}^{\infty}[\operatorname{Re}F(i\omega)\cos\omega t-\operatorname{Im}F(i\omega)\sin\omega t]\,d\omega, \qquad (3.49)$$

где Re и Im означают символы вещественной и мнимой частей соответственно. Последняя ступень преобразования законна в силу равенства (3.17). Как показывает соотношение



Фиг. 25

(3.49), импульс, приложенный к входу системы, возбуждает в ней все частоты в равной мере. Зная реакцию $h(t)$ системы на единичный импульс, можно найти ее частотную характеристику с помощью интеграла

$$F(i\omega)=\int_{0}^{\infty}h(t)\,e^{-i\omega t}\,dt. \qquad (3.50)$$

Для отдельных значений частот последний интеграл можно подсчитывать с помощью числовых методов.

Однако в практических условиях трудно осуществить входное воздействие в виде точного импульса. Более практичной формой входного воздействия служат единичный прямоугольный импульс и единичный треугольный импульс, изображенные на фиг. 25. Такие импульсы уже не будут возбуждать в одинаковой мере все частоты. Но если сделать длину τ импульса достаточно малой, то можно приблизиться к идеальному равномерному возбуждению по всем частотам.

Сименс и его сотрудники применили такой способ импульсного возбуждения для нахождения частотной характеристики самолета[1]) и разработали приближенный метод подсчета $F(i\omega)$ по закону изменения выходной координаты $y(t)$, найденному опытным путем. Кэрфмен и Гардинер[2]) обобщили этот метод обработки экспериментальных данных на случай входных функций любого вида.

3.6. Составление системы из элементов. Системы, рассмотренные в пп. 3.1, 3.3 и 3.4, в действительности представляют собой лишь элементы, входящие в состав гораздо более сложных систем, которые, как правило, приходится применять в технике устойчивых следящих систем.

Обратимся, например, к системе управления движением самолета по крену. Сигнал, управляющий перемещением элеронов, обычно поступает в форме электрического тока. Этот сигнал является входным для блока усилительных и счетных устройств[3]), представляющего собой довольно сложную электрическую цепь, которая может содержать электронные лампы. Процессы в блоке усилителя и счетных устройств определяются его передаточной функцией $F_1(s)$. Выход этого блока, в свою очередь, служит входом гидравлического исполнительного двигателя, приводящего элероны в движение. Процессы в гидравлическом двигателе определяются его передаточной функцией $F_2(s)$. Наконец, выходная переменная гидравлического двигателя — координата элеронов — служит входной величиной системы, характеризующей динамику самолета[4]). Динамика самолета дает передаточную функ-

[1]) Seamans R. C., Blasingame B. P., Clementson G. C., Journ. Aeronaut. Sci., **17**, 22 (1950).

[2]) Curfman H. J., Gardiner R. A., NACA TR 984 (1950).

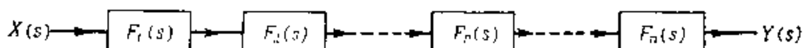
[3]) В простейшем случае счетное устройство сводится к устройству формирования управляющего сигнала η, зависящего от рассогласования ε в системе, например в виде четырехчленного аргумента управления $\eta = a\varepsilon + b\dot{\varepsilon} + c\ddot{\varepsilon} + d\int^{t} \varepsilon(\tau)\,d\tau$, где a, b, c, d — постоянные или функции времени. — *Прим. перев.*

[4]) *По крену. — Прим. перев.*

цию $F_3(s)$. Выходом системы, характеризующей динамику самолета, является движение самолета по крену. В этой задаче мы встречаемся с последовательным соединением различных элементов полной системы, начиная от команды по крену и кончая движением крена. Обозначим эту команду (входной сигнал) по крену через $x(t)$, а угол крена самолета — через $\varphi(t)$. Изображения переменных по Лапласу связаны, в силу сказанного выше, соотношением

$$\Phi_i(s) = F_3(s)\,F_2(s)\,F_1(s)\,X(s).$$

Поэтому передаточная функция всей системы управления движением самолета по крену равна произведению $F_1(s)\,F_2(s)F_3(s)$. Этот пример также иллюстрирует то обстоятельство, что передаточная функция в общем случае не



Фиг. 26

является безразмерной величиной, ибо она представляет собой отношение двух переменных различной размерности. Вход, т. е. команда по крену, представляет собой электрическое напряжение; выход же, т. е. угол крена самолета, имеет размерность угловой величины.

Таким образом, в общем случае, если система состоит из отдельных элементов с передаточными функциями $F_1(s)$, $F_2(s)$, ..., $F_r(s)$, ..., $F_n(s)$ с коэффициентами усиления K_1, K_2, ..., K_r, ..., K_n, причем эти элементы соединены последовательно (фиг. 26), передаточная функция $F(s)$ всей системы находится как произведение

$$F(s) = F_1(s)\,F_2(s)\ldots F_r(s)\ldots F_n(s), \tag{3.51}$$

а коэффициент K усиления всей системы равен произведению

$$K = K_1K_2\ldots K_r\ldots K_n. \tag{3.52}$$

Из равенства (3.51) видно, что передаточная функция $F(s)$ системы обладает всеми нулями и полюсами передаточных функций составляющих элементов. Эти нули и полюсы совместно с коэффициентом усиления K, подсчи-

тываемым по формуле (3.52), полностью определяют передаточную функцию $F(s)$.

Частотную характеристику системы можно записать в виде $F(i\omega)=Me^{i\theta}$. Если частотная характеристика r-го элемента имеет вид $M_r e^{i\theta_r}$, то, согласно (3.51),

$$Me^{i\theta}=(M_1e^{i\theta_1})(M_2e^{i\theta_2})\ldots(M_re^{i\theta_r})\ldots(M_ne^{i\theta_n}).$$

Поэтому

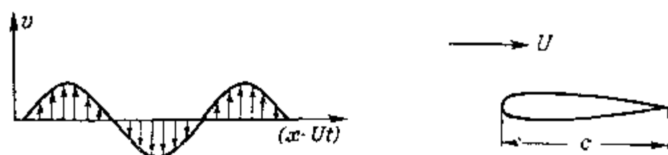
$$\left.\begin{aligned}\lg M=\lg M_1+\lg M_2+\ldots+\\+\lg M_r+\ldots+\lg M_n\\ \text{и}\qquad\qquad\qquad\\ \theta=\theta_1+\theta_2+\ldots+\theta_r+\ldots+\theta_n.\end{aligned}\right\}\qquad(3.53)$$

Соотношения (3.53) выявляют причину, в силу которой для построения диаграмм Боде выбрана логарифмическая шкала. Этот выбор шкалы облегчает работу по построению частотных характеристик системы, сводящуюся в этом случае к простому суммированию ординат характеристик составных элементов системы.

3.7. Трансцендентные передаточные функции. Метод преобразования Лапласа применим не только к задачам, определяемым обыкновенными линейными дифференциальными уравнениями с постоянными коэффициентами при заданных начальных условиях, но также и к линейным дифференциальным уравнениям в частных производных[1] с коэффициентами, не зависящими от времени t, и при граничных условиях, включающих начальные условия по t. В результате применения преобразования Лапласа к исходному дифференциальному уравнению в частных производных в преобразованном уравнении зависимость от переменной t заменяется зависимостью от параметра s. Это преобразованное уравнение представляет собой линейное дифференциальное уравнение по отношению к оставшимся независимым переменным и может быть решено при оставшихся граничных условиях. Конечно, методика, связанная с решением такого уравнения,

[1]) См., например, Churchill R. V., Modern operational methods in engineering, New York, 1944.

является гораздо более сложной, чем методика решения обыкновенных дифференциальных уравнений, рассмотренная в гл. II[1]). Но, с другой стороны, если рассматривать две какие-либо величины в решении преобразованного уравнения, то между ними существует линейная зависимость. Рассматривая одно из них как вход, а другое — как выход, заметим, что отношение выхода ко входу



Фиг. 27

в этом случае продолжает оставаться функцией параметра s, которую можно принять за передаточную функцию $F(s)$. Здесь существует, однако, одно различие: передаточная функция уже не выражается отношением двух многочленов по степеням s. Она является в общем случае трансцендентной функцией от s[2]).

Изучая движение крыла в двумерном потоке[3]), Сирс исследовал влияние небольших вертикальных порывов ветра скорости v на обтекание аэродинамического профиля хорды c горизонтальным потоком, текущим с постоянной скоростью U. Пусть v меняется синусоидально относительно x (x - координата, отсчитываемая в горизонтальном направлении по потоку) и относительно t (фиг. 27), так что

$$v(x, t) = \alpha_m U e^{i\omega\left(t - \frac{x}{U}\right)}. \qquad (3.54)$$

Здесь α_m — амплитуда, а ω — частота. Сирс показал, что при таком законе изменения скорости возмущающих

[1]) Если уравнение в частных производных зависит, помимо t, еще только от одной независимой переменной (одномерная задача), то преобразованное уравнение будет обыкновенным линейным дифференциальным уравнением относительно s. — *Прим. перев.*

[2]) К трансцендентным передаточным функциям приводят, в частности, задачи о системах, содержащих запаздывающие звенья. См. гл. VIII, а также [13, 29]. — *Прим. перев.*

[3]) Sears W. R., Journ. Aeronaut. Sci., 8, 104 (1941).

порывов ветра коэффициент C_l подъемной силы профиля (т. е. среднее значение подъемной силы профиля, приходящееся на единицу его площади и отнесенное к скоростному напору $\frac{1}{2}\rho U^2$, где ρ плотность воздуха, обтекающего профиль) определяется соотношением

$$C_l = 2\pi\alpha_m e^{i\omega t}\varphi(k), \tag{3.55}$$

причем

$$k = \frac{\omega c}{2U} \tag{3.56}$$

и

$$\varphi(k) = \frac{J_0(k)K_1(ik) + iJ_1(k)K_0(ik)}{K_1(ik) + K_0(ik)}. \tag{3.57}$$

Функции J_0, J_1 и K_0, K_1 в соотношении (3.57) суть функции Бесселя первого рода и приведенные функции Бесселя второго рода соответственно. Поэтому, рассматривая изображение $X(s)$ входа как изображение входной функции $v(0, t)$ и изображение $Y(s)$ выхода как изображение выходной функции $C_l(t)$, мы получим для передаточной функции $F(s)$ выражение

$$\frac{Y(s)}{X(s)} = F(s),$$

а для частотной характеристики $F(i\omega)$ — выражение

$$F(i\omega) = \frac{2\pi}{U}\varphi(k). \tag{3.58}$$

Таким образом, частотная характеристика представляет собой трансцендентную функцию.

Применение таких понятий, как трансцендентная передаточная функция и частотная характеристика, к задаче флаттера самолета было дано Дж. Дугунджи[1]).

[1]) Dugundji J., Journ. Aeronaut. Sci., 19, 422 (1952).

Глава IV

СИСТЕМЫ С ОБРАТНОЙ СВЯЗЬЮ

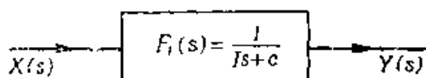
В этой главе мы введем центральное понятие современной техники автоматического регулирования и стабилизации: понятие обратной связи. Мы введем это понятие, начав с рассмотрения простейших систем — линейных систем с постоянными коэффициентами. Мы покажем, каким образом с помощью обратной связи достигается значительное повышение степени точности процессов регулирования и быстроты реакции системы на входной сигнал. Далее мы изложим принципы построения автоматических систем с обратной связью, обеспечивающие их устойчивость и оптимальную отработку.

4.1. Понятие обратной связи. Рассмотрим задачу о регулировании угловой скорости турбогенераторного агрегата. Основная цель регулирования заключается в поддержании значения угловой скорости в достаточной близости к ее заданной величине. Самый элементарный подход к решению этой задачи состоит в использовании так называемой разомкнутой системы регулирования, в которой мы стремимся уравновесить вращающий момент, развиваемый паровой турбиной, и момент, потребляемый генератором, т. е. момент нагрузки. Такого уравновешивания можно было бы достичь путем замера момента нагрузки с последующим дросселированием пара в соответствии с величиной этого момента. Однако есть все основания считать, что такой метод уравновешивания не может быть совершенным: всегда будет оставаться некоторый неуравновешенный момент (ошибка по моменту), который мы будем обозначать через $x(t)$. Этот остаточный момент будет сообщать ротору турбогенераторной установки некоторое угловое ускорение. Обозначив через $y(t)$ отклоне-

ние угловой скорости этого ротора от ее требуемого значения, через I — момент инерции ротора и через c — коэффициент демпфирования, обусловленный рассеиванием энергии, выпишем дифференциальное уравнение движения турбогенератора[1]:

$$I\frac{dy}{dt}+cy=x(t). \tag{4.1}$$

Блок-схема рассматриваемой нами системы представлена на фиг. 28. Легко видеть, что это уже известная нам система первого порядка, изученная в предыдущей главе, с постоянной времени, равной I/c; отношение же



Фиг. 28

установившегося значения отклонения скорости к остаточному моменту (ошибке по моменту) равно $1/c$. Но так как вес ротора турбогенераторной установки очень велик, то I представляет собой очень большую величину. Значение же коэффициента c демпфирования очень мало, так как потери в обмотке генератора невелики. Следовательно, постоянная времени системы чрезвычайно велика. А это означает, что всякое отклонение угловой скорости от ее нужного значения будет практически иметь место в течение длительного времени и устранение отклонений будет связано с трудностями. Кроме того, при малых отклонениях угловой скорости остаточный неуравновешенный момент будет неизбежно очень малым вследствие большой величины коэффициента $1/c$. Совершенно оче-

---

[1]) Это уравнение справедливо, если момент I инерции турбогенератора, т. е. ротора турбогенераторной установки, не меняется при изменении угловой координаты φ ротора. При $I=I(\varphi)$ надо исходить из уравнения

$$I(\varphi)\,d\dot{\varphi}/dt+\frac{1}{2}\dot{\varphi}^2I'(\varphi)=M_\varphi,$$

где M_φ — сумма моментов внешних сил, включая демпфирование.— *Прим. перев.*

видно, что такая система стабилизации угловой скорости турбогенератора практически бесполезна.

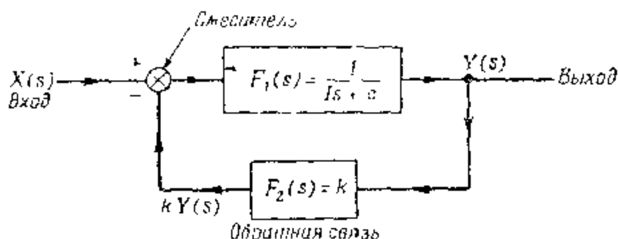
Выясним теперь, как изменится качество стабилизации угловой скорости при переходе к так называемой замкнутой системе регулирования. В замкнутой системе мы заставляем управляющий вращающий момент, с помощью которого выравнивается угловая скорость турбогенератора, зависеть от стабилизируемой переменной. Это означает переход к такой схеме регулирования, в которой открытие заслонки (дросселя), регулирующей поступление пара на лопатки турбины, зависит не только от нагрузки, но также и от величины y отклонения угловой скорости. Пусть эта вторая составляющая управляющего момента пропорциональна величине y с коэффициентом пропорциональности k. Когда угловая скорость превосходит ее стабилизируемое значение ($y > 0$), то заслонка прикрывается и вращающий момент, вызывающий ускоренное вращение ротора, уменьшается на величину ky. Когда угловая скорость падает по сравнению с этим значением, вращающий момент увеличивается на величину ky. Таким образом, в замкнутой системе движение ротора описывается дифференциальным уравнением

$$I \frac{dy}{dt} + (c + k)\, y = x(t). \tag{4.2}$$

Единственное различие между уравнениями (4.1) и (4.2) заключается в замене коэффициента c в (4.1) суммой $c + k$ в (4.2). Следовательно, в этих условиях постоянная времени системы равна $I/(c+k)$, а отношение установившегося отклонения угловой скорости к остаточному моменту равно $1/(c+k)$. Поэтому переход к замкнутой системе регулирования позволяет в значительной мере уменьшить как постоянную времени, так и ошибку по угловой скорости путем выбора параметра k гораздо бо́льшим, чем c, причем это очень легко осуществить, так как c мало. Благодаря этому имеется возможность разрабатывать замкнутые системы регулирования, предназначенные для точного и быстрого регулирования и тем самым для большого повышения качества переходных процессов.

Блок-схему замкнутой системы можно представить в виде схемы, показанной на фиг. 29; при этом переда-

точная функция самого турбогенератора как объекта регулирования сохраняется в том же виде, в каком она изображена на фиг. 28. На фиг. 29 мы ввели одно условное обозначение, принятое в схемах систем с обратной связью: на схеме указаны знаки сложения и вычитания сигналов, поступающих на вход суммирующего звена (смесителя). Если пересечение двух линий помечено



Фиг. 29

только точкой, то в этом пересечении не происходит ни сложения, ни вычитания. Здесь данная величина только «измеряется». Так, отклонение y угловой скорости измеряется на выходе турбогенератора и используется для образования управляющего вращающего момента. Как следует из рассмотрения фиг. 29, замкнутая система характеризуется наличием цепи обратной связи. Поэтому всю систему в целом естественно называть системой автоматического регулирования с обратной связью.

В разобранном выше простом примере преимущество замкнутой системы над системой разомкнутой можно выявить путем сравнения дифференциальных уравнений (4.1) и (4.2). В более же сложных системах удобно провести исследование лишь с привлечением метода передаточной функции. Этот метод излагается в следующих разделах.

4.2. Критерии для разработки систем с обратной связью. Рассмотрим общий случай системы автоматического регулирования с обратной связью и с передаточными функциями $F_1(s)$ и $F_2(s)$, подобными соответствующим передаточным функциям в схеме на фиг. 29. $F_1(s)$ называется передаточной функцией прямой цепи, а $F_2(s)$ —

передаточной функцией цепи обратной связи. В этих условиях изображение $X(s)$ входа связано с изображением $Y(s)$ выхода соотношением

$$Y(s) = F_1(s)[X(s) - F_2(s)Y(s)].$$

Разрешив последнее соотношение относительно $Y(s)$, получим

$$\frac{Y(s)}{X(s)} = \frac{F_1(s)}{1 + F_1(s)F_2(s)} = F_s(s), \tag{4.3}$$

где $F_s(s)$ представляет собой, таким образом, передаточную функцию системы или, иначе, отношение изображения выхода всей системы к ее входу.

Для удобства дальнейших рассуждений в выражениях для передаточных функций целесообразно выделить в явном виде коэффициенты усиления K_1 и K_2 в передаточных функциях $F_1(s)$ и $F_2(s)$ соответственно. Таким образом, мы запишем

$$\left.\begin{aligned} F_1(s) &= K_1G_1(s), \\ F_2(s) &= K_2G_2(s). \end{aligned}\right\} \tag{4.4}$$

Очевидно, что функция $G(s)$ представляет собой безразмерную величину; размерность же передаточной функции целиком вошла в коэффициент усиления K. Функция $G(s)$ заключает в себе все «сведения» о строении передаточной функции; иными словами, $G(s)$ определяет нули и полюсы передаточной функции. Во всех дальнейших рассуждениях мы будем обычно рассматривать влияние передаточной функции на процессы в системе как результат двух отдельных воздействий: влияния расположения нулей и полюсов передаточной функции, т. е. влияния функции $G(s)$, и влияния величины коэффициента усиления K. Это разделение влияний величин K и $G(s)$ оправдывается еще и тем обстоятельством, что строение передаточной функции определяется строением счетного устройства блока усилительных и счетных звеньев внутри полной системы, характеризуемой функцией $F(s)$, а величина коэффициента усиления определяется усилительным устройством этого блока. Более того, при проектировании всей замкнутой системы автоматического регулирования на величины K и $G(s)$ можно влиять почти независимо. В силу этих обстоятельств

коэффициент K и функцию $G(s)$ можно изменять, по существу, порозь и рассматривать их также порозь.

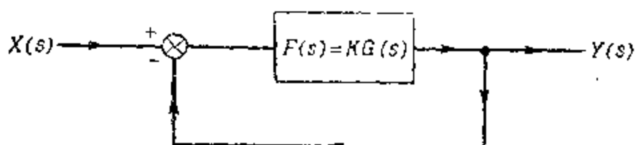
Используя соотношения (4.4), перепишем передаточную функцию $F_s(s)$ в виде

$$\frac{Y(s)}{X(s)} = F_s(s) = \frac{K_1G_1(s)}{1+K_1G_1(s)K_2G_2(s)} = \frac{1}{\frac{1}{K_1G_1(s)}+K_2G_2(s)}. \qquad (4.5)$$

Обозначим через $E(s)$ изображение по Лапласу рассогласования (ошибки) $e(t)$ (3.11); тогда

$$\frac{E(s)}{Y(s)} = \frac{X(s)-Y(s)}{Y(s)} = \frac{1}{F_s(s)} - 1 = \frac{1}{K_1G_1(s)} - [1-K_2G_2(s)]. \qquad (4.6)$$

В системах с простой обратной связью, подобных системе, блок-схема которой изображена на фиг. 30,



Фиг. 30

передаточная функция $F_2(s)$ цепи обратной связи сводится попросту к единице. Это означает, что для использования в системе управления с обратной связью выходная переменная только измеряется, но не подвергается преобразованию. В этом случае выражения (4.5) и (4.6) передаточных функций упрощаются:

$$\frac{Y(s)}{X(s)} = F_s(s) = \frac{KG(s)}{1+KG(s)} = \frac{1}{\frac{1}{KG(s)}+1} \qquad (4.7)$$

$$\frac{E(s)}{Y(s)} = \frac{1}{KG(s)}. \qquad (4.8)$$

Первым требованием, предъявляемым к системе автоматического регулирования, или следящей системе, является требование устойчивости. Это означает, что выходное движение $y(t)$ системы нужно демпфировать, за исключением, быть может, тех случаев, когда выход должен

изменяться в установившемся режиме по синусоидальному закону. Возвращаясь к исследованию, проведенному нами в п. 2.4, отметим следующее: в математическом отношении условие устойчивости равносильно утверждению о том, что функция $F_s(s)$ не должна иметь полюсов в правой полуплоскости s, в которой вещественная часть s положительна. В случае системы с простой обратной связью из соотношения (4.8) следует, что полюсами функции $F_s(s)$ являются нули функции

$$\frac{1}{F_s(s)} = \frac{1}{K_1 G_1(s)} + K_2 G_2(s). \tag{4.9}$$

В системах с простой обратной связью полюсы функции $F_s(s)$, в силу (4.7), служат нулями функции

$$\frac{1}{F_s(s)} = \frac{1}{KG(s)} + 1. \tag{4.10}$$

Отсюда мы получаем первый критерий, которому необходимо удовлетворять при проектировании систем с обратной связью:

а) *Функция* $1/F_s(s)$, *определяемая соотношениями (4.9) и (4.10), не должна иметь нулей в правой полуплоскости комплексной переменной* s.

Вторым требованием, предъявляемым к системам автоматического регулирования, или следящим системам, является быстрота реакции (отработки). Пусть s_r — полюс функции $F_s(s)$. Тогда, вспоминая результаты исследования, проведенного в п. 2.4, отметим, что выход системы содержит слагаемое, пропорциональное $e^{s_r t}$. Таким образом, быстрота реакции определяется величиной числа s_r. Чем больше величина s_r, тем короче шкала времени и, следовательно, тем быстрее протекает реакция системы. В результате мы приходим ко второму критерию проектирования систем с обратной связью, формулируемому следующим образом:

б) *Все нули функции* $1/F_s(s)$, *определяемые соотношениями* (4.9) *и* (4.10), *должны быть в достаточной степени удалены влево от мнимой оси на плоскости комплексной переменной* s[1]).

---

[1]) Более строго этот критерий сформулирован Я. З. Цыпкиным и П. В. Бромбергом [30], впервые его предложившими.—*Прим. перев.*

Если система с обратной связью предназначена для отслеживания на выходе входного сигнала, то выходная система должна (в установившемся режиме, т. е. по исчезновении влияния переходного процесса) как можно точнее воспроизводить течение входной переменной системы. Поэтому для таких позиционных систем имеет место третье требование, заключающееся в том, что отношение $E(0)/Y(0)$ изображения установившейся ошибки к установившемуся выходу должно быть возможно меньшим. Это требование можно выразить в виде условия, налагаемого на коэффициенты усиления передаточных функций (4.6) и (4.8).

Таким образом, мы приходим к третьему критерию:

в) *В позиционных следящих системах для достижения точности отработки в общем случае систем с обратной связью* (4.5) *требуется выполнение условия*

$$\frac{1}{K_1} - (1 - K_2) \sim 0 \tag{4.11}$$

и в случае систем с простой обратной связью (4.7)—*условия*

$$K \gg 1. \tag{4.12}$$

Требования (а), (б) и (в) представляют собой критерии, которым необходимо удовлетворять при проектировании следящих систем. На практике обычно бывает трудно полностью удовлетворить условиям (б) и (в) в желаемой мере, и в общем случае приходится прибегать к компромиссу, что мы рассмотрим подробнее в следующих разделах.

4.3. Метод Найквиста. Как уже отмечалось выше, передаточные функции обычно представляют собой отношения двух многочленов по степеням s. Поэтому критерий (а), сформулированный в предыдущем разделе, вообще говоря, равносилен требованию отсутствия у многочлена корней, имеющих вещественную положительную часть. Эту классическую задачу решил Э. Дж. Раус, получив так называемые неравенства Рауса, составленные по определенному способу с помощью коэффициентов данного

многочлена[1]). Однако инженеры, работающие в области регулирования, неохотно пользуются этим критерием из-за неясности характера изменения левых частей неравенств Рауса при изменении коэффициентов многочлена[2]). Инженеры предпочитают метод исследования, основанный на непосредственном использовании передаточных функций, определяемых соотношениями (4.9) и (4.10), без дополнительных преобразований. Это предпочтение обусловливается тем, что упомянутые передаточные функции непосредственно отражают все имеющиеся данные о рассматриваемой системе и инженеры их ощущают «физически».

Такой способ был разработан Г. Найквистом. Способ Найквиста основан на следующей теореме Коши об аналитической функции $f(s)$ комплексной переменной s[3]).

Если функция $f(s)$ имеет n нулей и m полюсов внутри области, ограниченной замкнутым контуром C, то при одном обходе этого контура по направлению движения часовой стрелки вектор $f(s)$ совершит $n-m$ оборотов вокруг начала координат также по направлению движения часовой стрелки.

Для применения этой весьма сильной теоремы к нашей задаче выберем контур C так, чтобы он охватывал всю правую часть плоскости комплексной переменной s, в которой будут расположены нули с вещественной положительной частью. Такой контур изображен на фиг. 31; он образован мнимой осью и полуокружностью радиуса $R \to \infty$, лежащей вправо от этой оси. Рассмотрим сперва более простой случай, а именно, систему с простой обратной связью. С помощью соотношения (4.10) мы установили, что полюсы функции $1/F_s(s)$ суть нули функции $G(s)$. Пусть функция $G(s)$ имеет m нулей в правой полуплоскости комплексной переменной s. Тогда функция

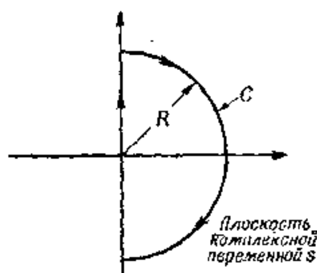
---

[1]) Эти неравенства, называемые обычно неравенствами Рауса—Гурвица, приводятся почти в любом курсе теоретической механики и теории устойчивости и регулирования. Доказательство необходимости и достаточности этих неравенств имеется в [31, 32] (индуктивный метод) и в [3] (дедуктивный метод). — *Прим. перев.*

[2]) И тем более при изменении первичных параметров системы. — *Прим. перев.*

[3]) См., например, Уиттекер, Ватсон, Курс современного анализа, М.—Л., 1933, 6.31.

$1/F_s(s)$ имеет внутри контура C m полюсов. Поэтому для того, чтобы функция $1/F_s(s)$ не имела нулей в правой полуплоскости комплексной переменной, необходимо, чтобы вектор $1/F_s(s)$ совершил m оборотов вокруг начала в направлении против вращения часовой стрелки при обходе контура C, изображенного на фиг. 31, при $R \to \infty$. Как легко видеть, обращаясь к соотношению (4.10), это требование равносильно требованию о том, чтобы вектор $1/KG(s)$ совершил m оборотов против направления движения часовой стрелки вокруг точки -1[1]). Но так как коэффициент усиления K представляет собой постоянную, последнее требование равносильно тому, чтобы вектор $1/G(s)$ совершил m оборотов против направления движения часовой стрелки вокруг точки $-K$. Само собой разумеется, что в тех случаях, когда функция $G(s)$ вовсе не имеет нулей в правой полуплоскости комплексной переменной s, т. е. при $m = 0$, критерий устойчивости, сформулированный Найквистом, требует, чтобы вектор $1/G(s)$ не совершал вокруг точки $-K$ ни одного оборота.



Фиг. 31

Проиллюстрируем применение этого способа на примере простой передаточной функции:

$$G(s) = \frac{1}{s(1+\tau_1 s)(1+\tau_2 s)}\,.$$

Тогда

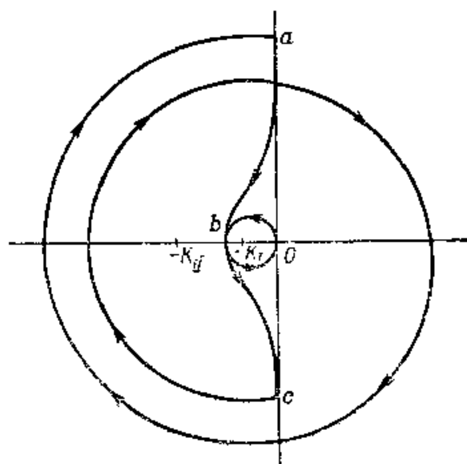
$$\frac{1}{G(s)} = s(1+\tau_1 s)(1+\tau_2 s). \qquad (4.13)$$

Рассмотрим прежде всего ту часть контура C, изображенного на фиг. 31, которая совпадает с мнимой осью; для этой части контура C

$$\frac{1}{G(i\omega)} = i\omega(1+i\tau_1\omega)(1+i\tau_2\omega).$$

[1] См. [9, 13, 20, 31]. — *Прим. перев.*

При $\omega = 0$, $1/G(i\omega) = i0$. Когда $\omega \to \infty$, $1/G(i\omega) \to -i\infty$. Поэтому с возрастанием ω от 0 до ∞ абсолютная величина вектора $1/G(i\omega)$ возрастает, а его фазовый угол увеличивается от $\pi/2$ до $3\pi/2$. При отрицательных значениях ω траектория конца вектора $1/G(i\omega)$ является просто отражением соответствующей траектории для положительных ω относительно вещественной оси, что следует из



Фиг. 32

свойства, выражаемого равенством (3.17). Таким образом, когда s пробегает вдоль мнимой оси, конец вектора $1/G(i\omega)$ описывает кривую $abOc$, показанную на фиг. 32.

Когда же точка s пробегает вдоль окружности достаточно большого радиуса, изображенной на фиг. 31, то $1/G(s) \sim s^3$. Следовательно, по мере перемещения точки s в направлении вращения часовой стрелки от $i\infty$ до $-i\infty$ вектор $1/G(s)$ будет поворачиваться также в этом направлении, но в три раза быстрее. Этот участок траектории конца вектора $1/G(s)$, таким образом, определяется кривой, простирающейся от c до a на фиг. 32. Рассматривая траектории, нанесенные на этой фигуре, мы убедимся, что если придать коэффициенту усиления K значение K_{I}, указанное на фигуре, то вектор $1/G(s)$ в конечном счете не совершит ни одного оборота вокруг точки $-K_{\text{I}}$.

Так как функция $G(s)$, определяемая равенством (4.13), не имеет нулей, то это обстоятельство означает, что данная система с обратной связью устойчива. Если же $K=K_{II}$, то вектор $1/G(s)$ совершит около точки $-K_{II}$ два полных оборота. Поэтому при этом большем значении K следящая система с обратной связью будет неустойчива. Действительно, при этом значении K передаточная функция $F_s(s)$ имеет два полюса с положительной вещественной частью. Граничной точкой, соответствующей переходу от устойчивости к неустойчивости, является точка b (фиг. 32). Значения K, соответствующие устойчивой системе, должны лежать между началом и этой точкой.

В общем случае следящей системы задача исследования устойчивости этой системы сводится к выяснению вопроса о том, имеет ли функция $1/F_s(s)$, определяемая соотношением (4.9), хотя бы один нуль в правой полуплоскости комплексной переменной s. Непосредственное применение теоремы Коши к правой части (4.9) связано с определенными неудобствами, так как при этом приходится складывать два вектора: $1/K_1G_1(s)$ и $K_2G_2(s)$. Обозначим через m_1 и m_2 число нулей в правой полуплоскости комплексной переменной s функций $G_1(s)$ и $G_2(s)$ соответственно, а через n_1 и n_2 — число полюсов функций $G_1(s)$ и $G_2(s)$ в этой же области. Очевидно, что число полюсов функции $1/F_s(s)$ в правой полуплоскости s равно m_1+n_2. Разделим теперь $1/F_s(s)$ на $K_2G_2(s)$; частное, полученное в результате этого деления, будет иметь m_2 полюсов и n_2 нулей. Но может оказаться, что некоторые из нулей совпадут с некоторыми из полюсов и в этом случае соответствующие нули и полюсы сократятся. Обозначим через α число исчезнувших таким образом нулей и полюсов. Упомянутое же частное равно

$$\frac{1}{F_s(s)\,K_2G_2(s)}=\frac{1}{K_1K_2G_1(s)\,G_2(s)}+1. \qquad (4.14)$$

Число полюсов функции $1/F_s(s)\,K_2G_2(s)$ в правой полуплоскости комплексной переменной s равно числу полюсов функции $1/K_1K_2G_1(s)\,G_2(s)$ и равно, таким образом, m_1+m_2. Предположим теперь, что функция $1/F_s(s)$ не имеет нулей в правой полуплоскости комплексной переменной s, т. е. что система с обратной связью устойчива.

Для этого случая число нулей и полюсов различных функций указано в табл. 2. Отсюда легко сделать вывод, что $n_2 - \alpha = 0$ и что функция $1/F_s(s)\,K_2G_2(s)$ также не имеет нулей в правой полуплоскости s.

Таблица 2

Функция	Число (в правой полуплоскости комплексной переменной s)	
	нулей	полюсов
$G_1(s)$	m_1	n_1
$G_2(s)$	m_2	n_2
$1/F_s(s)$	0	$m_1 + n_2$
$1/F_s(s)\,K_2G_2(s)$	$n_2 - \alpha$	$m_1 + n_2 + m_2 - \alpha = m_1 + m_2$

Следовательно, при обходе контура C, изображенного на фиг. 31, вектор $1/F_s(s)\,K_2G_2(s)$ должен совершить $-(m_1 + m_2)$ оборотов по часовой стрелке вокруг начала. Обращаясь же к равенству (4.14), установим, что это условие устойчивости равносильно следующему: вектор $1/K_1K_2G_1(s)\,G_2(s)$ должен совершить, при обходе контура C, $m_1 + m_2$ оборотов против часовой стрелки около точки -1. Иначе говоря, вектор $1/G_1(s)\,G_2(s)$ должен совершить $m_1 + m_2$ оборотов против часовой стрелки около точки $-(K_1K_2)$. В этом и заключается критерий Найквиста для устойчивости системы автоматического регулирования (системы с обратной связью) общего типа.

Существенной частью контура интегрирования, используемого в теории Найквиста, служит мнимая ось, где $s = i\omega$, что наглядно показано в только что разобранном примере (фиг. 32). Поэтому для исследования вопроса об устойчивости системы можно непосредственно исходить из частотных характеристик прямой цепи и цепи обратной связи. Так как частотные характеристики составных звеньев системы часто находятся экспериментальным путем, то метод, позволяющий непосредственно использовать экспериментальные данные, обладает определенным преимуществом. Именно в этом заключается достоинство способа Найквиста. Недостаток же этого способа связан с неясностью

в отношении степени (запаса) устойчивости при пользовании им. Иными словами, если система устойчива, то в какой мере она демпфирована? Для ответа на этот вопрос мы можем изменить рассмотренный критерий и потребовать, чтобы функция $1/F_s(s)$ не имела нулей правее некоторой прямой, параллельной мнимой оси на плоскости комплексной переменной s и расположенной влево от этой оси[1]). Минимальная величина демпфирования определяется расстоянием $-\lambda$ между такой прямой и мнимой осью. При этом можно опять пользоваться критерием Найквиста, должным образом изменив контур C, вдоль которого перемещается точка s, и количество нулей m, m_1 и m_2, содержащихся в соответствующих передаточных функциях. Однако для фактического применения критерия Найквиста в данной задаче необходимо располагать значениями передаточных функций не при $s = i\omega$, а при $s = -\lambda + i\omega$. А в этих условиях уже нельзя непосредственно использовать частотные характеристики звеньев системы и, таким образом, способ Найквиста теряет здесь свою главную ценность. В этом отношении гораздо лучше метод, разработанный У. Р. Ивэнсом[2]), который мы рассмотрим в следующем пункте.

4.4. Метод Ивэнса. Начнем наше исследование со случая системы автоматического регулирования с простой обратной связью. В такой системе основная задача сводится к нахождению корней уравнения

$$0 = \frac{1}{F_s(s)} = 1 + \frac{1}{KG(s)} \tag{4.15}$$

при заданной функции $G(s)$. Метод Ивэнса заключается в определении этих корней в функции коэффициента усиления K и называется поэтому методом геометрического места корней. Коль скоро найдена эта зависимость корней от K, то всякое сочетание корней определяет соответствующее значение K. Таким образом, этот метод приводит к гораздо более широким результатам по сравнению с обеспечением только требования (а) п. 4.2 и доставляет

---

[1]) См. примечание на стр. 67.—*Прим. ред.*

[2]) Evans W. R., Transaction of the American Institute of Electrical Engineers, 67, 547—551 (1948).

фактическое решение задачи проектирования систем с обратной связью на основе всех трех требований, выдвинутых в этом пункте.

Определим строение функции $G(s)$ с помощью ее нулей $p_1, p_2, \ldots, p_m$ и полюсов $q_1, q_2, \ldots, q_n$. Учитывая определение коэффициента усиления, доставляемое равенствами (3.16), (3.21) и (3.23), напишем

$$G(s) = A\frac{(s-p_1)(s-p_2)\ \ldots\ (s-p_m)}{(s-q_1)(s-q_2)\ \ldots\ (s-q_n)}, \tag{4.16}$$

где

$$A = \frac{(-q_1)(-q_2)\ \ldots\ (-q_n)}{(-p_1)(-p_2)\ \ldots\ (-p_m)}.$$

Для реальных систем коэффициенты многочленов, входящих в числитель и знаменатель выражения для функции $G(s)$, суть вещественные числа. Поэтому для таких систем нули $p_1, \ldots, p_m$ являются вещественными числами или образуют комплексные сопряженные пары чисел. Подобным же образом и полюсы $q_1, \ldots, q_n$ представляют собой вещественные или попарно сопряженные комплексные числа. Значит, и коэффициент A всегда будет вещественным. В прикладных задачах исходные параметры или переменные системы определяются таким образом, чтобы вещественная величина A была, кроме того, и положительной. Поэтому в дальнейшем мы будем считать A вещественным положительным числом. В общем случае степень знаменателя функции $G(s)$ равна степени числителя или превосходит ее, т. е. $n \geqslant m$.

Выразим каждый сомножитель числителя и знаменателя правой части (4.16) в векторной форме:

$$\left.\begin{aligned} s-p_1 &= P_1e^{i\varphi_1},\\ s-p_2 &= P_2e^{i\varphi_2},\\ &\ldots\ldots\ldots\\ s-p_m &= P_me^{i\varphi_m}, \end{aligned}\right\} \tag{4.17}$$

$$\left.\begin{aligned} s-q_1 &= Q_1e^{i\theta_1},\\ s-q_2 &= Q_2e^{i\theta_2},\\ &\ldots\ldots\ldots\\ s-q_n &= Q_ne^{i\theta_n}. \end{aligned}\right\} \tag{4.18}$$

вектор $P_re^{i\varphi_r}$ имеет началом точку p_r, а концом — точку s.

Вектор $Q_r e^{i\theta_r}$ имеет началом точку q_r, а концом — точку s (s — текущая точка на плоскости комплексной переменной s). Используя равенства (4.17) и (4.18), запишем $G(s)$ в виде

$$G(s) = A\,\frac{(P_1 e^{i\varphi_1})\,(P_2 e^{i\varphi_2})\,\ldots\,(P_m e^{i\varphi_m})}{(Q_1 e^{i\theta_1})\,(Q_2 e^{i\theta_2})\,\ldots\,(Q_n e^{i\theta_n})}. \qquad (4.19)$$

Так как A - вещественное положительное число, то можно написать вместо (4.19)

$$G(s) = Re^{i\Theta}, \qquad (4.20)$$

где

$$R = A\,\frac{(P_1 P_2\,\ldots\,P_m)}{(Q_1 Q_2\,\ldots\,Q_n)} \qquad (4.21)$$

и

$$\Theta = (\varphi_1 + \varphi_2 + \ldots + \varphi_m) - (\theta_1 + \theta_2 + \ldots + \theta_n). \qquad (4.22)$$

Так как величины $P_1, \ldots, P_m$ и $Q_1, \ldots, Q_n$ суть модули векторов, определяемых равенствами (4.17) и (4.18), то они положительны. Поэтому положительной является и величина R. С помощью новых величин основное уравнение, определяющее корни обратной передаточной функции системы (4.15), преобразуется к виду

$$\frac{e^{-i\Theta}}{KR} = -1.$$

Поэтому, чтобы удовлетворить этому уравнению, необходимо обеспечить выполнение соотношений

$$KR = 1 \qquad (4.23)$$

и

$$\Theta = \pm\pi. \qquad (4.24)$$

Метод Ивэнса содержит два этапа: первый этап заключается в определении всех значений независимой переменной s, удовлетворяющих условию (4.24), наложенному на аргумент комплексной переменной $G(s)$. Далее, зная такое геометрическое место корней, мы можем подсчитать с помощью (4.23) R и тем самым K вдоль всего этого геометрического места. Ивэнс разработал несколько полезных правил для построения геометрического места корней, к изложению которых мы и перейдем.

Правило 1. Так как, в силу уравнения (4.15), при $K=0$ $G(s)\to\infty$, то нули функции $1/F_c(s)$ являются полюсами функции $G(s)$. Это означает, что геометрическое место корней начинается в полюсах функции $G(s)$. Полюсы $G(s)$ мы будем отмечать точками на плоскости комплексной переменной s.

Правило 2. Так как при $K\to\infty$ $G(s)\to 0$, то при $K\to\infty$ геометрическое место корней проходит через нули функции $G(s)$. Эти нули мы будем отмечать маленькими кружками на плоскости комплексной переменной s. При $n>m$ число нулей функции $G(s)$ меньше числа нулей функции $1/F_c(s)$. Но в этом случае $G(s)\to 0$ при $s\to\infty$. Поэтому недостающим нулям соответствует значение $s=\infty$. Далее, при достаточно больших s

$$G(s)\sim\frac{A}{s^{n-m}}.$$

Поэтому при таких значениях s уравнение (4.15) можно заменить приближенным уравнением

$$s^{n-m}\approx -KA.$$

Таким образом определяются фазовые углы асимптот геометрического места корней:

$$\frac{\pi}{n-m}+\frac{2k\pi}{n-m},\qquad k=1,\ 2,\ 3,\ \ldots \tag{4.25}$$

Правило 3. На вещественной оси геометрическое место корней представляет собой последовательность сегментов, соединяющих нули и полюсы функции $G(s)$, расположенные на этой оси. Началом геометрического места корней на вещественной оси служит нуль, расположенный правее всех остальных.

В справедливости этого правила легко убедиться, рассматривая произвольную точку s на вещественной оси: углы между направлениями, проведенными в эту точку из двух сопряженных нулей или полюсов, и вещественной осью равны $+\varphi$ и $-\varphi$ или $+\theta$ и $-\theta$ соответственно. Таким образом, сумма каждой из этих пар углов равна нулю. Угол же между направлением, проведенным в эту точку из какого-либо нуля или полюса на вещественной оси, и самой вещественной осью равен нулю для всех полюсов или нулей, лежащих влево от рассматриваемой

точки s, и равен π для всех нулей или полюсов, лежащих вправо от s. Таким образом, соответствующая сумма равна π, если вправо от точки находится нечетное число полюсов и нулей функции $G(s)$.

Правило 4. Если геометрическое место корней отклоняется от вещественной оси, то положение точки, в которой геометрическое место отходит от этой оси, можно оценить из того условия, что при малом отклонении $\Delta\omega$ от вещественной оси приращение угла[1]), обусловленное влиянием полюсов и нулей функции $G(s)$, расположенных на оси влево от искомой точки, должно уничтожаться приращением этого же угла, обусловленным влиянием полюсов и нулей функции $G(s)$, расположенных вправо от этой точки.

Пример. Рассмотрим передаточную функцию

$$G(s)=\frac{(0{,}001)(2)(6)}{(s+0{,}001)(s+2)(s+6)}. \tag{4.26}$$

При $K=0$ геометрическое место корней исходит из точек $-0{,}001$, -2 и -6, лежащих на вещественной оси. Отрезки геометрического места лежат между точками $-0{,}001$ и -2 и между -6 и $-\infty$. Здесь $m=0$, $n=3$. Поэтому, в силу (4.25), фазовые углы асимптот равны $+\pi/3$, $-\pi/3$ и π. Применяя правило 4, получаем

$$\frac{\Delta\omega}{\lambda_1+0{,}001}+\frac{\Delta\omega}{\lambda_1+2}+\frac{\Delta\omega}{\lambda_1+6}=0$$

или

$$(\lambda_1+2)(\lambda_1+6)+(\lambda_1+0{,}001)(\lambda_1+6)+(\lambda_1+0{,}001)(\lambda_1+2)=0.$$

Отсюда λ_1 определяется квадратным уравнением

$$3\lambda_1^2+16{,}002\lambda_1+12{,}008=0,$$

т. е.

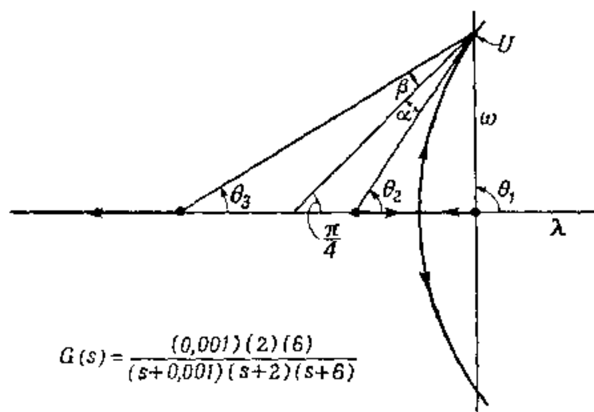
$$\lambda_1=-\frac{16{,}002}{6}-\sqrt{\left(\frac{16{,}002}{6}\right)^2-\frac{12{,}008}{3}}=-0{,}904.$$

Правило 5. Положение точки, в которой геометрическое место корней пересекает мнимую ось при пере-

[1]) Имеется в виду угол, определенный в правиле 3.— *Прим. перев.*

ходе в правую полуплоскость комплексной переменной s, часто можно оценить, пренебрегая малым по абсолютной величине полюсом функции $G(s)$.

Пример. Рассмотрим опять передаточную функцию (4.26). На значительном удалении от начала ее можно



Фиг. 33

с очень хорошей точностью аппроксимировать функцией

$$G(s) \approx \frac{(0{,}001)(2)(6)}{s(s+2)(s+6)}.$$

Тогда $\theta_1 \approx \pi/2$ (фиг. 33) и, следовательно, условие (4.24) сводится к равенству

$$\theta = -\pi = -\theta_2 - \theta_3 - \frac{\pi}{2}$$

или

$$\theta_2 + \theta_3 \approx \frac{\pi}{2}.$$

Но, рассматривая график на фиг. 33, мы замечаем, что

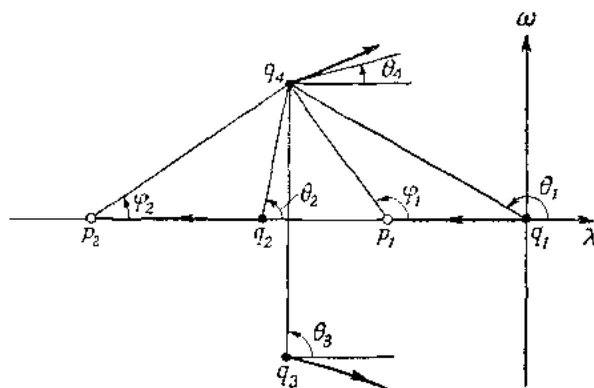
$$\frac{\pi}{4} = \theta_3 + \beta, \qquad \frac{\pi}{4} + \alpha = \theta_2,$$

откуда следует

$$\alpha \approx \beta.$$

Последнее равенство и представляет собой условие для определения точки U пересечения.

Правило 6. Направление касательной к геометрическому месту корней при выходе его из какого-либо полюса (или при подходе его к какому-либо нулю) нетрудно определить путем подсчета угла между этой касательной в данном полюсе (или нуле) и вещественной осью.



Фиг. 34

При таком подсчете используются зависимости между фазовыми углами для всех нулей и полюсов[1]), расположенных, по условию, в левой полуплоскости комплексной переменной s.

Пример. На фиг. 34 изображено геометрическое место корней передаточной функции $G(s)$, имеющей два нуля и два полюса на вещественной оси и одну пару комплексных сопряженных полюсов. При достаточно малом удалении точки s от полюса q_1 углы φ_1, φ_2, θ_1, θ_2 и θ_3, соответствующие остальным нулям и полюсам, останутся неизменными. Таким образом, в силу (4.24), угол θ_4 найдется из уравнения

$$(\varphi_1+\varphi_2)-[(\theta_1+\theta_2+\theta_3)+\theta_4]=\pi.$$

Перечисленные правила определяют основные свойства геометрического места корней. Это геометрическое место вне нулей и полюсов функции $G(s)$ находится с помощью построения по точкам, после чего можно определить ха-

[1]) См. (4.24). — *Прим. перев.*

рактер изменения K вдоль построенной таким образом кривой. После того как желаемое расположение корней уравнения $1/F_з(s)=0$ выбрано, находится соответствующее значение K.

Тем самым завершается процесс синтеза системы с обратной связью[1]).

4.5. Гидродинамическая аналогия метода геометрического места корней. Сочетая соотношения (4.15) и (4.16), получаем

$$\frac{(s-q_1)(s-q_2)\ldots(s-q_n)}{(s-p_1)(s-p_2)\ldots(s-p_m)}=-KA.$$

Логарифмируя обе части этого равенства и деля результат на 2π, приходим к соотношению

$$W(s)=\frac{1}{2\pi}\sum_{i=1}^{n}\ln(s-q_i)-\frac{1}{2\pi}\sum_{j=1}^{m}\ln(s-p_j)=$$
$$=\frac{1}{2\pi}\ln KA+i\left(\frac{1}{2}\right). \qquad (4.27)$$

Соотношение (4.27) можно физически истолковать различными способами. Чрезвычайно наглядный способ истолкования состоит в том, чтобы рассматривать функцию $W(s)$ в качестве комплексного потенциала скоростей плоского безвихревого течения идеальной несжимаемой жидкости[2])[3]). Обозначим через $\varphi(\lambda, \omega)$ и $\psi(\lambda, \omega)$ потенциальную функцию и функцию тока данного течения соответственно. Тогда

$$W(s)=\varphi(\lambda, \omega)+i\psi(\lambda, \omega), \qquad (4.28)$$

где $s=\lambda+i\omega$. В таком случае соотношение (4.27), рассматриваемое как уравнение геометрического места кор-

---

[1]) Следует указать, что процессом синтеза, вообще говоря, называется значительно более сложная операция, см. [5, 9, 16]. — *Прим. ред.*

[2]) См., например, Streeter V. L., Fluid dynamics, New York, 1948.

[3]) См. также [34, 35]. — *Прим. перев.*

ней для передаточной функции $1/F_s(s)$, можно истолковать, как уравнение линий, вдоль которых функция тока ψ сохраняет постоянное значение, равное $1/2$. Пользуясь языком гидромеханики, можно сказать, что геометрическое место корней состоит, таким образом, из ветвей линий тока, соответствующих постоянной, равной $1/2$. Что касается потенциальной функции, то она испытывает изменение вдоль линий тока и равна

$$\frac{1}{2\pi} \ln KA.$$

Соотношение (4.27) означает, что данное течение возникло под действием n источников единичной интенсивности, расположенных в точках $q_1, q_2, \ldots, q_n$, и m стоков, также единичной интенсивности, расположенных в точках $p_1, p_2, \ldots, p_m$. На наших графиках, иллюстрирующих распределение нулей и полюсов на плоскости комплексной переменной s, источники помечены точками, а стоки — кружочками. Такое истолкование сразу позволяет «физически ощутить» смысл геометрического места корней, примеры которого графически изображены на фиг. 33 и 34.

Гидродинамическая аналогия очень полезна также и для выявления изменений, которые целесообразно внести в цепь обратной связи системы для повышения качества системы. Пусть, например, система определяется передаточной функцией

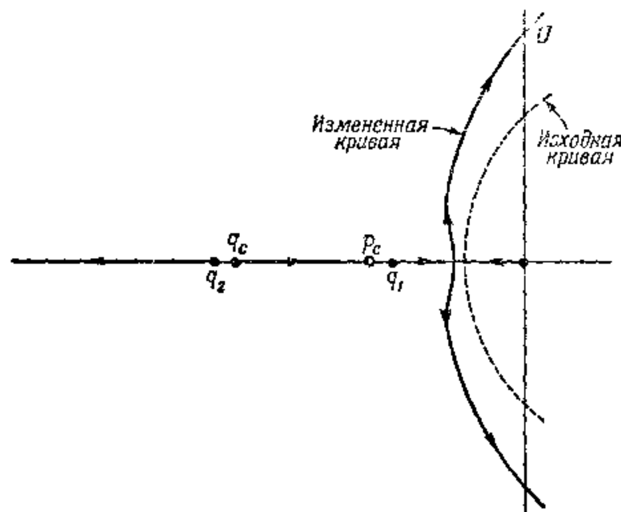
$$G(s) = \frac{q_1 q_2}{s(s-q_1)(s-q_2)}, \qquad |q_1| < |q_2|.$$

Недостаток этой системы заключается в том, что она неустойчива в замкнутом режиме при весьма низких значениях коэффициента усиления K и потому не может удовлетворить требованию (в) п. 4.2. Геометрическое место корней для этой системы подобно геометрическому месту, изображенному на фиг. 33. Прибегая к гидродинамической аналогии, мы немедленно обнаружим, что точку U пересечения геометрического места с мнимой осью можно переместить вверх, сдвинув влево часть линии тока в окрестности точки U; для этого достаточно ввести сток p_c вблизи полюса q_1 и источник q_c вблизи полюса q_2. Измененная таким образом переда-

точная функция будет иметь вид

$$G(s)=\frac{q_c}{p_c}\cdot\frac{(s-p_c)}{(s-q_c)}\,\frac{q_1q_2}{s(s-q_1)(s-q_2)},$$

а соответствующее геометрическое место корней изображено на фиг. 35. Так как $|p_c|<|q_c|$, то дополнительная передаточная функция, которая на блок-схеме измененной



Фиг. 35

системы подключается последовательно к исходной передаточной функции, должна определять фазоопережающий контур, подобный контуру, рассмотренному в п. 3.3 и характеризуемому передаточной функцией (3.27).

Гидродинамическая аналогия способствует также выявлению возможностей повышения быстродействия системы, медленно реагирующей на входное воздействие, путем внесения необходимых изменений в цепь обратной связи. В силу критерия (б) п. 4.2 для быстрой отработки системы необходимо, чтобы вещественная часть корней характеристического уравнения была достаточно большой по абсолютной величине. Рассмотрим для простоты линейную механическую систему первого порядка, характеризуемую малым по абсолютной величине полюсом q_1 передаточной функции, лежащим на вещественной оси. Если

к такой системе последовательно подключить сильно демпфированный электрический контур, характеризуемый большим по абсолютной величине полюсом q_2 его передаточной функции, также лежащим на отрицательной вещественной полуоси, то качество переходного процесса на выходе системы не улучшится, так как в ней попрежнему останется малый по абсолютной величине полюс q_1. Замкнем в этом случае систему с помощью цепи обратной связи. Обращаясь к линиям тока, т. е. к геометрическому месту корней, мы замечаем, что с увеличением коэффициента усиления K меньший по абсолютной величине корень q_1 характеристического уравнения смещается влево в направлении корня q_2, большего по абсолютной величине. Поэтому путем подходящего выбора коэффициента K можно увеличить абсолютную величину этих корней так, чтобы они лежали гораздо левее первоначального положения корня q_1, и тем самым резко увеличить быстроту реакции системы.

Метод геометрического места корней можно применить также и к системам автоматического регулирования с обратной связью общего вида. В этом случае дело сводится к построению геометрического места корней уравнения $1/F_k(s)=0$, где функция $1/F_k(s)=1+1/KG(s)$ определяется соотношением (4.15). Таким образом, геометрическое место корней определяется условием

$$\frac{1}{K_1G_1(s)}=-K_2G_2(s).$$

Так как корни уравнения (4.15) отличны от нулей функции $G_2(s)$, то можно разделить обе части последнего равенства на $G_2(s)/K_1$, в результате чего будем иметь

$$\frac{1}{G_1(s)\,G_2(s)}=-K_1K_2. \tag{4.29}$$

Введя обозначения

$$\begin{gathered} G(s)=G_1(s)\,G_2(s), \\ K=K_1K_2 \end{gathered} \tag{4.30}$$

и сравнив уравнения (4.29) и (4.15), мы обнаружим, что задача построения геометрического места корней для системы автоматического регулирования с произвольной

структурой цепи обратной связи свелась к аналогичной задаче для системы автоматического регулирования с простой обратной связью, т. е. к только что разобранной задаче. Действительно, выполненное в п. 4.3 подробное исследование, касающееся применения способа Найквиста к системе автоматического регулирования с обратной связью общего вида, показывает, что метод приведения, определяемый соотношениями (4.30), пригоден и в этой задаче. Следовательно, с учетом соотношений приведения (4.30) исчезает какое-либо различие между системами автоматического регулирования с простой обратной связью и системами с обратной связью общего вида в той мере, в какой это касается обеспечения качества регулирования, характеризуемого критериями (а), (б) и (в), установленными в разделе 4.2. И только в задачах, связанных с обеспечением определенных количественных закономерностей в системах с обратной связью, необходимо должным образом учитывать различие между строением передаточных функций $F_s(s)$, характеризуемых выражениями (4.3), с одной стороны, и (4.7) — с другой.

4.6. Метод Боде. В точке U на фиг. 33 и 35, в которой геометрическое место корней пересекает мнимую ось при переходе в правую полуплоскость комплексной переменной s, корень, по определению, является чисто мнимой величиной; обозначим ее через $i\omega^*$. Другими словами, уравнение (4.15) удовлетворяется при $s = i\omega^*$, т. е.

$$KG(i\omega^*) = F(i\omega^*) = -1 = 1 \cdot e^{-i\pi}.$$

Поэтому критическое состояние системы в смысле перехода от устойчивости к неустойчивости имеет место при том условии, что амплитуда M частотной характеристики равна единице и в то же время фазовый угол θ этой характеристики равен $-\pi$. Это условие для критического состояния можно вывести также и из критерия Найквиста, согласно которому критическое состояние системы определяется прохождением характеристики $1/F(i\omega)$ через точку -1[1]). В самом деле, обращаясь к типовому примеру

[1]) По критерию Рауса — Гурвица это критическое состояние системы определяется обращением в нуль предпоследнего определителя Гурвица. — *Прим. перев.*

частотной характеристики, каким является характеристика на фиг. 32, мы заключаем, что для устойчивости системы эта характеристика должна охватывать точку -1. Так как абсолютная величина функции $1/F(i\omega)$ в общем случае возрастает с ростом ω, то охват этой характеристикой точки -1 можно обеспечить, сделав абсолютную величину функции $1/F(i\omega)$ большей единицы при том значении ω, при котором фазовый угол (аргумент) комплексной величины $1/F(i\omega)$ равен π. Последнее требование равносильно требованию о том, чтобы число M было меньше единицы при $\theta = -\pi$. Иначе говоря, величина θ должна быть больше, чем $-\pi$, при $M = 1$. Это условие устойчивости лежит в основе метода Боде; значение частоты, при которой амплитуда M частотной характеристики равна единице, называется частотой среза, а разность между θ и $-\pi$ называется избытком фазы (запасом по фазе). В этих терминах критерий устойчивости Боде требует, чтобы при частоте среза избыток фазы составлял от 30° до 50°. На логарифмической амплитудно-частотной характеристике частота среза равна частоте, при которой $\lg M = 0$, и это обстоятельство позволяет легко применять критерий Боде.

Метод Боде аналогичен методу Найквиста в том отношении, что он допускает непосредственное использование частотных характеристик. Однако наряду с этим преимуществом метода Боде, связанным с его простотой, последний обладает по сравнению с методом Ивэнса тем недостатком, что с его помощью не представляется возможным определить степень устойчивости. Р. М. Осборн[1]) предпринял попытку улучшить в этом отношении метод Боде и получил полуэмпирическую формулу для подсчета коэффициента ζ демпфирования, соответствующего наибольшему критическому корню:

$$\zeta \approx \frac{1}{60} \frac{\alpha}{m}. \tag{4.31}$$

Здесь α — избыток фазы при частоте среза, m — наклон логарифмической амплитудно-частотной характеристики

[1]) Осборн Р. М., Доклад, прочитанный на летней сессии Института инженеров-радистов (IRE), Сан-Франциско, август 1949.

($\lg M$ в функции ω) при частоте среза. При этом единица измерения времени для ζ та же, что и для ω. Так, если $\alpha = 30°$, а $m = 1{,}7$, то $\zeta \approx 1/(2 \times 1{,}7) = 0{,}3$.

4.7. Вопросы синтеза передаточной функции. Различные методы исследования устойчивости систем автоматического регулирования (систем с обратной связью), рассмотренные в предыдущих пунктах, являются главным образом методами анализа. Частично они представляют собой и методы синтеза, т. е. методы построения передаточной функции, но лишь в той мере, в которой они предоставляют возможность установить диапазон возможных значений коэффициента усиления K. Конечно, указания о возможных изменениях строения передаточной функции, позволяющих улучшить качество переходных процессов в системе, можно получить как с помощью метода анализа, так и с помощью метода синтеза. Это в особенности относится к методу геометрического места корней. Однако осуществление выясненных таким образом желаемых изменений в передаточной функции путем соответственных изменений физических звеньев системы зависит главным образом от инженерного искусства в области систем автоматического регулирования.

Общее решение этой задачи синтеза известно только в одном отношении — в отношении синтеза электрической цепи, состоящей из активных сопротивлений и емкостей, т. е. цепи типа RC, по заданным нулям и полюсам передаточной функции такой цепи. Так как подобные цепи обладают большой гибкостью и очень часто используются для компенсации передаточных функций других элементов в системе и так как требуемые изменения передаточной функции часто можно свести к заданию дополнительных нулей и полюсов, то наличие этого общего решения задачи синтеза передаточной функции представляется весьма важным. В решение этой задачи важный вклад внесли Э. А. Гильемен[1]) и Л. Вейнберг[2])[3]). Мы не будем здесь

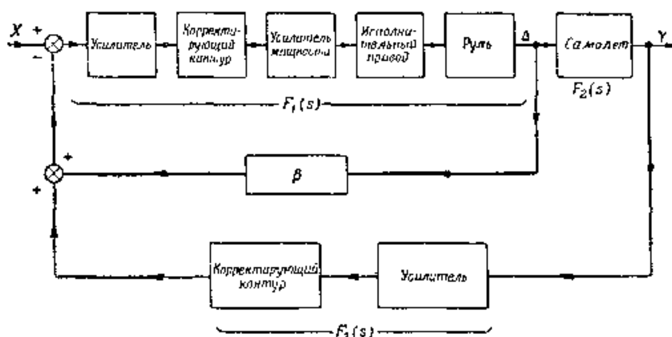
[1]) Guillemin E. A., Journ. Math. and Phys., **28**, 22—44 (1949).

[2]) Weinberg L., Journ. Appl. Phys., **24**, 207—216 (1953).

[3]) Подробное изложение этих работ имеется в сборнике [32], гл. V.—*Прим. перев.*

углубляться в рассмотрение данного вопроса, а только подчеркнем имеющуюся возможность физически осуществлять цепи типа RC, обладающие очень сложными свойствами.

4.8. Многоконтурные системы автоматического регулирования. До сих пор мы изучали только одноконтурные системы автоматического регулирования. В инженерной практике часто приходится прибегать к более сложным системам. Так, например, на фиг. 36 изображена



Фиг. 36

блок-схема типичной системы[1]) автоматического управления вращательным движением самолета около одной из осей. Внутренняя цепь представляет собой так называемую внутреннюю обратную связь по движению руля или жесткую обратную связь. При разомкнутой внутренней цепи мы возвращаемся к обычной системе управления с обратной связью, так что имеет место соотношение

$$\frac{Y(s)}{X(s)} = \frac{F_1(s)\,F_2(s)}{1 + F_1(s)\,F_2(s)\,F_3(s)}\,. \tag{4.32}$$

Если же замкнуты обе цепи, то

$$\Delta(s) = F_1(s)\,[X(s) - \beta\Delta(s) - F_3(s)\,Y(s)]$$

$$Y(s) = F_2(s)\,\Delta(s),$$

[1]) Becker L., Aeronaut. Eng. Rev., September 1951, p. 17.

Поэтому

$$\frac{\Delta(s)}{X(s)}=\frac{F_1(s)}{1+\beta F_1(s)+F_1(s)F_2(s)F_3(s)} \tag{4.33}$$

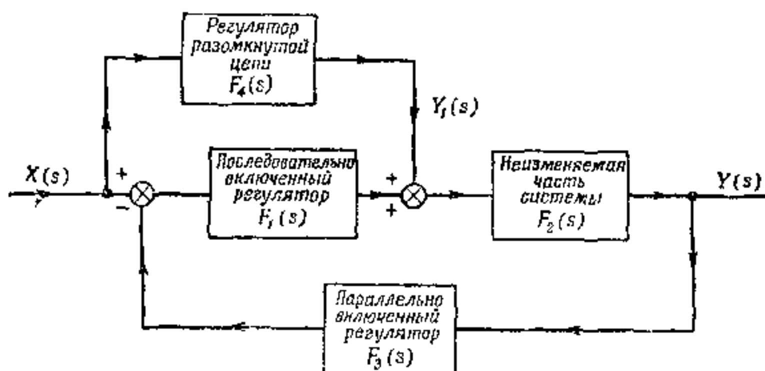
и

$$\frac{Y(s)}{X(s)}=\frac{F_1(s)F_2(s)}{1+\beta F_1(s)+F_1(s)F_2(s)F_3(s)}\,. \tag{4.34}$$

Следовательно, устойчивость и характер реакции системы управления зависят от значений нулей функции

$$1+\beta F_1(s)+F_1(s)F_2(s)F(s).$$

Одно из затруднений, возникающих при разработке хорошей системы управления, заключается в необходимости одновременного обеспечения точности управления



Фиг. 37

и тем самым высокого коэффициента усиления K и достаточного быстродействия вместе с удовлетворительным демпфированием. Это обстоятельство привело к идее сочетания систем замкнутого цикла с системами разомкнутого цикла, предложенной Дж. Р. Муром [1]) [2]). Для уяснения этой идеи рассмотрим систему, схематически представленную на фиг. 37; в этой системе осуществлено параллельное соединение замкнутой системы и разомкнутой системы.

[1]) Moore J. R., Proc. IRE, 39, 1421—1432 (1951).

[2]) См. также [36].

Исходя из схемы фиг. 37, получим

$$\left.\begin{aligned} F_4(s)X(s) &= Y_1(s) \\ \text{и} \quad Y(s) &= F_2(s)\{Y_1(s) + F_1(s)[X(s) - F_3(s)Y(s)]\}. \end{aligned}\right\} \quad (4.35)$$

Исключив $Y_1(s)$ из уравнений последней системы и разрешив полученное уравнение относительно $Y(s)$, будем иметь

$$\frac{Y(s)}{X(s)} = \frac{F_2(s)F_4(s) + F_1(s)F_2(s)}{1 + F_1(s)F_2(s)F_3(s)}. \quad (4.36)$$

Таким образом, устойчивость и быстрота реакции системы зависят от нулей функции $1 + F_1(s)F_2(s)F_3(s)$.

Так как функция $F_2(s)$ фиксирована, то задача синтеза заключается в разыскании подходящих передаточных функций $F_1(s)$ и $F_3(s)$. Действительное течение выхода системы, в особенности величина установившейся ошибки, зависит от передаточной функции $F_4(s)$ регулятора разомкнутой цепи. Поэтому цепь обратной связи предназначена главным образом для обеспечения устойчивости системы и должного качества переходных процессов, тогда как установившиеся (или «синхронизирующие») значения переменных определяются в значительной мере действием разомкнутой части системы.

Если система предназначена для одновременного регулирования нескольких переменных, которые к тому же связаны между собой, как это имеет место, например, в паровой машине, то схема системы состоит из большого числа замкнутых контуров со сложными линиями обратных связей [1]). Примером наиболее сложной системы такого типа служит, вероятно, система автоматического управления самолетом [2]). Хотя исследование такой системы и подчиняется тем же принципам, на основании которых выполняется исследование более простых систем автоматического регулирования, изложенное в этой главе, однако анализ столь сложной системы практически невыполним без использования электронных моделей. Но в этом и заключается единственный путь технического развития, состоящего в процессе перехода от принципов к практическому осуществлению.

---

[1]) См., например, Hänny J., Regelungstheorie, Zürich, 1946.

[2]) Rea J., Aeronaut. Eng. Rev., November 1951, p. 139.

Глава V

АВТОНОМНОЕ РЕГУЛИРОВАНИЕ

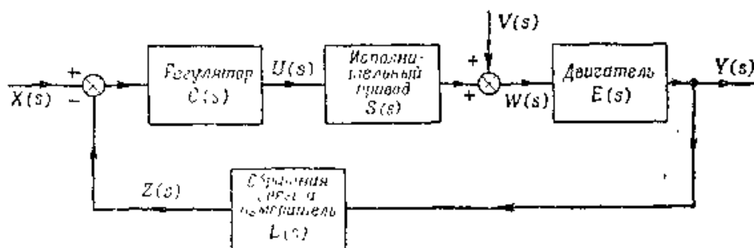
Для сложных систем с несколькими регулируемыми переменными и при наличии взаимодействия между этими регулируемыми переменными в общем случае требуется добавить новый критерий, которому такая система должна удовлетворять: критерий автономности[1]). Например, состояние турбореактивного двигателя с дожиганием определяется следующими переменными: угловой скоростью турбокомпрессорного агрегата, количеством топлива, подаваемого в камеры горения в единицу времени, и площадью поперечного сечения выходного отверстия для газов в хвостовой части двигателя. Однако управление этим двигателем можно вести, исходя из фиксированных значений скорости, подачи топлива в камеры горения в единицу времени и подачи топлива в камеру дожигания в единицу времени. В таких условиях становится очевидным, что один из критериев, на основе которых ведется разработка системы автоматического управления двигателем, должен требовать независимости настройки по каждой из трех регулируемых переменных: изменение режима подачи топлива в камеру дожигания не должно отражаться на угловой скорости ротора двигателя, а изменение этой угловой скорости не должно требовать изменения режима питания топливом камер горения[2]). Решение этой конкретной задачи заключается в разработке

[1]) Автономность регулирования означает, что при возбуждении переходного процесса только по какой-либо одной из регулируемых переменных значения остальных регулируемых переменных не должны изменяться под влиянием этого процесса. — *Прим. перев.*

[2]) Очевидно, что последнее требование может выполняться лишь в определенном диапазоне скоростей, так как, например, при расходе топлива, минимально достаточном для режима малого газа, нельзя добиться большего числа оборотов никакими приемами регулирования. — *Прим. перев.*

определенного метода регулирования выходного сечения двигателя в соответствии со значениями других переменных и в разработке надлежащей системы автоматики. Цель настоящей главы и заключается в изложении общего метода синтеза такого устройства для автономного регулирования систем любой сложности. Этот общий метод впервые дали А. С. Боксенбом и Р. Худ[1][2]).

5.1. Регулирование системы с одной переменной. Рассмотрим прежде всего простую систему с единственным регулируемым выходом $y(t)$ и единственным входом $x(t)$, который можно настраивать на определенное значение.



Фиг. 38

Пусть $Y(s)$ и $X(s)$ суть изображения по Лапласу соответственных величин. Обратимся к системе регулирования, выполненной на основе схемы, показанной на фиг. 38. На этой схеме $E(s)$ передаточная функция двигателя, $L(s)$ — передаточная функция цепи обратной связи, $S(s)$ — передаточная функция исполнительного привода и $C(s)$ передаточная функция регулятора. Здесь проектировщик может легко изменять только функцию $C(s)$. Эта система незначительно отличается от простой системы, приведенной на фиг. 29, в том отношении, что между исполнительным приводом и двигателем введено возмущение $V(s)$, выражающее случайные внешние влияния.

---

[1]) Boksenbom A. S., Hood R., NACA TR 980 (1950).

[2]) Подобный метод был впервые предложен Вознесенским И. И., см. [37].—*Прим. ред.*

Изображение $W(s)$ входа двигателя связано с изображением $Y(s)$ его выхода с помощью соотношения

$$Y(s)=E(s)W(s)=E(s)[S(s)U(s)+V(s)], \qquad (5.1)$$

где $U(s)$ — изображение выхода регулятора, в свою очередь определяемое соотношением

$$U(s)=C(s)[X(s)-Z(s)]=C(s)[X(s)-L(s)Y(s)]. \quad (5.2)$$

Исключив $U(s)$ из соотношений (5.1) и (5.2), получим

$$Y(s)=\frac{E(s)S(s)C(s)}{E(s)S(s)C(s)L(s)+1}X(s)+ \\ +\frac{E(s)}{E(s)S(s)C(s)L(s)+1}V(s). \quad (5.3)$$

Это соотношение дает изображение выхода при определенных начальных значениях функции $y(t)$ и определенном входе $x(t)$. Если в выражении, стоящем в правой части равенства (5.3), отбросить второе слагаемое, характеризующее влияние возмущения $V(s)$, то это выражение совпадет с выражением (4.3), ранее найденным для простых систем автоматического регулирования, и исследование свойств рассматриваемой здесь системы можно провести подобным же образом. Однако в случае более сложных систем эту схему необходимо обобщить; к такому обобщению мы сейчас и перейдем.

5.2. Регулирование системы со многими переменными [1]. Пусть двигатель имеет i выходов

$$Y_1(s),\ Y_2(s),\ \ldots,\ Y_\nu(s),\ \ldots,\ Y_i(s)$$

и n входов

$$W_1(s),\ W_2(s),\ \ldots,\ W_k(s),\ \ldots,\ W_n(s).$$

---

[1]) Такие системы называются многомерными, см. [38]. — *Прим. перев.*

В этих условиях обобщением соотношения (5.1) служит система равенств

$$\left.\begin{aligned} Y_1(s) &= E_{11}(s)\,W_1(s) + \\ &\quad + E_{12}(s)\,W_2(s) + \cdots + E_{1n}(s)\,W_n(s), \\ Y_2(s) &= E_{21}(s)\,W_1(s) + \\ &\quad + E_{22}(s)\,W_2(s) + \cdots + E_{2n}(s)\,W_n(s), \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ Y_i(s) &= E_{i1}(s)\,W_1(s) + \\ &\quad + E_{i2}(s)\,W_2(s) + \cdots + E_{in}(s)\,W_n(s). \end{aligned}\right\} \tag{5.4}$$

Каждая из функций $E_{jk}(s)$ представляет собой передаточную функцию, которая при приложении входного воздействия $W_k(s)$ доставляет одну из составляющих изображения $Y_j(s)$ выхода системы по j-й координате. Следовательно, функция $E_{jk}(s)$, вообще говоря, представляет собой отношение двух многочленов относительно s; функция $E_{jk}(s)$ находится или аналитическим путем с помощью расчета двигателя, или экспериментально с помощью снятия его частотных характеристик. Систему соотношений (5.4) можно записать в более сжатом виде:

$$Y_\nu(s) = \sum_{k=1}^{n} E_{\nu k}(s)\,W_k(s). \tag{5.5}$$

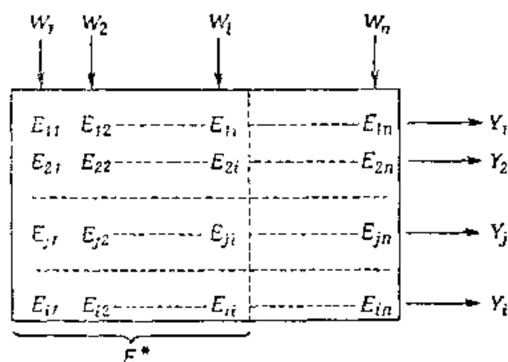
Таблица коэффициентов $E_{\nu k}(s)$ образует матрицу, которую можно в соответствии с ее смыслом назвать матрицей передаточных функций; условимся обозначать ее через E. Мы можем тогда считать, что входы $W_k(s)$ «входят» в соответствующие столбцы матрицы E, а выходы $Y_\nu(s)$ «выходят» из соответствующих строк этой матрицы[1]), согласно схеме на фиг. 39. Мы будем интересоваться тем

---

[1]) Систему (5.5) можно записать в матричном виде: $Y = EW$, где Y — матрица-столбец $\begin{Vmatrix} Y_1(s) \\ Y_2(s) \\ \cdots \\ Y_i(s) \end{Vmatrix}$, а W — матрица-столбец $\begin{Vmatrix} W_1(s) \\ W_2(s) \\ \cdots \\ W_n(s) \end{Vmatrix}$; см. также [13, 39, 40]. — *Прим. перев.*

случаем, когда число входов превышает число выходов или равно ему, т. е. $n \geqslant i$. Поэтому матрица E – прямоугольная, причем число столбцов превышает число строк [1]). В дальнейшем квадратную матрицу, образуемую элементами только первых i столбцов, мы будем обозначать через E^*.

Так как число входных координат двигателя превышает число выходных координат, то для полного определения



Фиг. 39

поведения системы, кроме значений $X_j(s)$ $(j = 1, \ldots, i)$ выходов $Y_\nu(s)$, требуется также задать значения $\Xi_\mu(s)$ $(\mu = i - 1, \ldots, n)$ переменных $W_\mu(s)$. Следовательно, регулируемыми переменными будут i выходов $Y_\nu(s)$ $(\nu = 1, \ldots, i)$ и $n - i$ входов $W_\mu(s)$ двигателя. Обозначим через $\Upsilon_\mu(s)$ значения $\hat{W}_\mu(s)$, измеренные на входе звена обратной связи. Тогда рассогласования системы выразятся разностями $\Xi_\mu(s) - \Upsilon_\mu(s)$. Рассогласования по выходным переменным двигателя определяются разностями $X_\nu(s) - Z_\nu(s)$, где $Z_\nu(s)$ $(\nu = 1, \ldots, i)$ представляют собой значения этих выходных переменных, измеренные на выходе звена обратной связи (на фиг. 38). Назначение регулятора состоит в том, чтобы воспринимать эти рассогласования как его входные переменные и вырабатывать сигналы $U_k(s)$, управляющие перемещением исполнительных приводов. В этом заключается цель введения обратной связи.

---

[1]) При $n > i$. — *Прим. перев.*

Обобщение системы управления на случай многомерной системы состоит в том, что управляющие сигналы составляются, как линейные формы от всех рассогласований. Так как имеется n сигналов рассогласований, то число управляющих сигналов также равно n; таким образом, k принимает все значения от 1 до n включительно, и мы получаем

$$\begin{aligned} U_1(s) &= C_{11}(X_1 - Z_1) + C_{12}(X_2 - Z_2) + \cdots + C_{1i}(X_i - Z_i) + \\ &\quad + C'_{1,\,i+1}(\Xi_{i+1} - \Upsilon_{i+1}) + \cdots + C'_{1n}(\Xi_n - \Upsilon_n), \\ U_2(s) &= C_{21}(X_1 - Z_1) + C_{22}(X_2 - Z_2) + \cdots + C_{2i}(X_i - Z_i) + \\ &\quad + C'_{2,\,i+1}(\Xi_{i+1} - \Upsilon_{i+1}) + \cdots + C'_{2n}(\Xi_n - \Upsilon_n), \quad (5.6) \\ &\dots\dots\dots\dots\dots\dots\dots\dots \\ U_n(s) &= C_{n1}(X_1 - Z_1) + C_{n2}(X_2 - Z_2) + \cdots + C_{ni}(X_i - Z_i) + \\ &\quad + C'_{n,\,i+1}(\Xi_{i+1} - \Upsilon_{i+1}) + \cdots + C'_{nn}(\Xi_n - \Upsilon_n), \end{aligned}$$

где мы разбили матрицу передаточных функций, с помощью которых образуются управляющие сигналы, на две матрицы C и C' для того, чтобы в явной форме выделить два вида рассогласований. Равенства (5.6) можно записать в более сжатом виде[1]):

$$U_k(s) = \sum_{\nu=1}^{i} C_{k\nu}(X_\nu - Z_\nu) + \sum_{\mu=i+1}^{n} C'_{k\mu}(\Xi_\mu - \Upsilon_\mu); \quad (5.7)$$

$$\nu = 1,\ 2, \cdots,\ i; \quad k = 1,\ 2, \cdots,\ n.$$

Каждая из величин $C_{k\nu}$ и $C'_{k\mu}$, конечно, представляет собой отношение двух многочленов относительно s. Соотношения (5.6) и (5.7) можно также изобразить графически с помощью схемы, приведенной на фиг. 40.

Величины $Z_\nu(s)$ и $\Upsilon_\mu(s)$, измеряемые на выходе блока обратной связи, выражаются через величины $Y_\nu(s)$ и $W_\mu(s)$

---

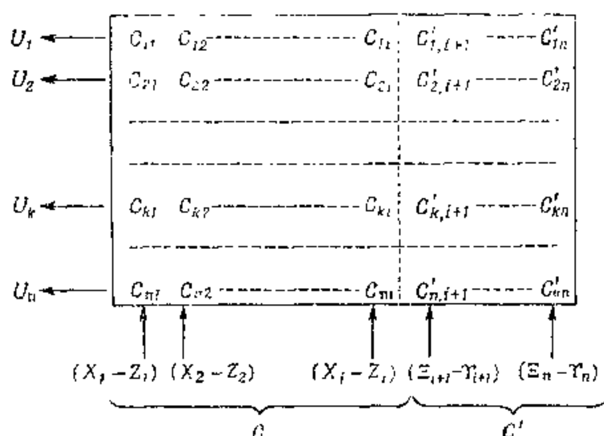
[1]) Или в матричной форме: $U = C(X - Z) + C'(\Xi - \Upsilon)$, где $U = \| U_k(s) \|\ (n,\ 1)$; $C = \| C_{k\nu} \|\ (n,\ i)$; $X = \| X_\nu \|\ (i,\ 1)$; $Z = \| Z_\nu \|\ (i,\ 1)$; $C' = \| C'_{k\mu} \|\ (n,\ n-i-1)$; $\Xi = \| \Xi_\mu \|\ (n-i-1,\ 1)$; $\Upsilon = \| \Upsilon_\mu \|\ (n-i-1,\ 1)$. — *Прим. перев.*

с помощью передаточных функций $L_{\nu\nu}(s)$ и $L_{\mu\mu}(s)$ этого блока:

$$Z_\nu(s) = L_{\nu\nu}(s)\,Y_\nu(s), \tag{5.8}$$

$$\Upsilon_\mu(s) = L_{\mu\mu}(s)\,W_\mu(s). \tag{5.9}$$

Каждый управляющий сигнал будет действовать на соответствующий исполнительный привод. Выходные же координаты этих приводов совместно с возмущениями



Фиг. 40

$V_k(s)$ определяют входные координаты $W_k(s)$ двигателя. Если мы обозначим через $S_{kh}(s)$ передаточные функции исполнительных приводов, то

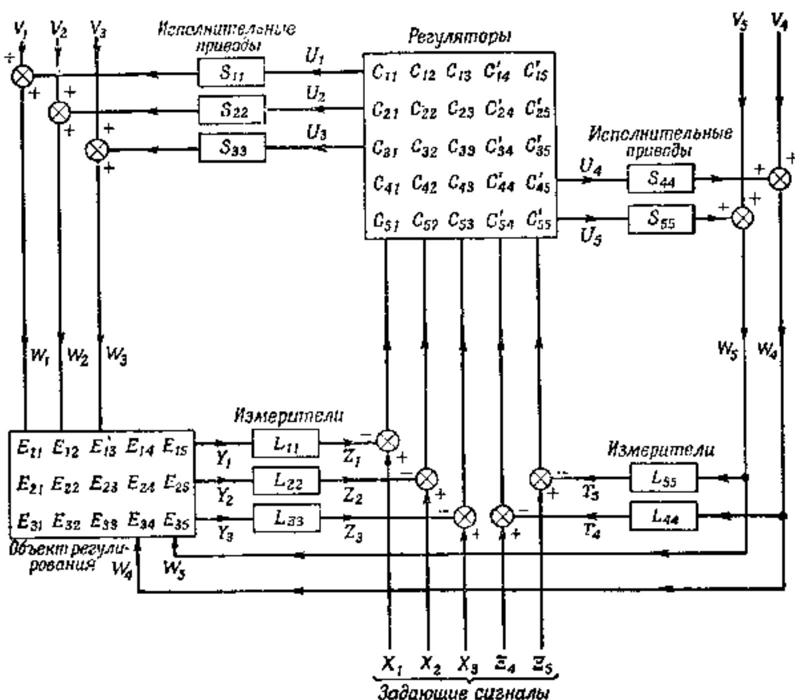
$$W_k(s) = S_{kh}(s)\,U_k(s) + V_k(s), \tag{5.10}$$
$$k = 1,\ 2,\ \ldots,\ n.$$

Соотношения от (5.4) до (5.10) полностью описывают действие системы управления двигателем со многими регулируемыми параметрами.

На фиг. 41 изображена блок-схема системы с тремя выходными координатами $Y_1(s)$, $Y_2(s)$ и $Y_3(s)$ двигателя и двумя регулируемыми входными воздействиями $W_4(s)$ и $W_5(s)$, приложенными к двигателю. Вся система оказывается замкнутой, за исключением задающих сигналов

и внешних возмущающих воздействий, которые могут быть приложены к системе извне.

Исключив функции $U_k(s)$, $Z_\nu(s)$ и $\Upsilon_\mu(s)$ из написанной выше системы уравнений, мы получим выражения для



Фиг. 41

$Y_j(s)$ и $W_k(s)$ в виде

$$Y_j(s) = \sum_{k=1}^{n} \left\{ \sum_{\nu=1}^{i} E_{jk}(s) S_{kk}(s) C_{k\nu}(s) [X_\nu(s) - L_{\nu\nu}(s) Y_\nu(s)] + \right.$$

$$+ \sum_{\mu=i+1}^{n} E_{jk}(s) S_{kk}(s) C'_{k\mu}(s) [\Xi_\mu(s) -$$

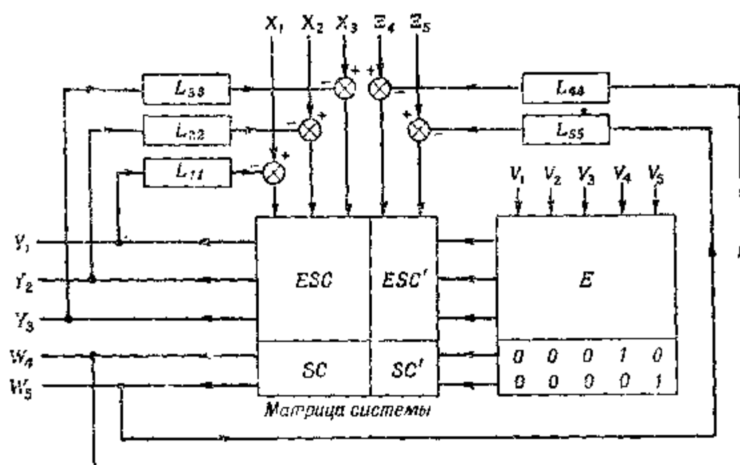
$$\left. - L_{\mu\mu}(s) W_\mu(s)] + E_{jk}(s) V_k(s) \right\} \qquad (5.11)$$

и

$$W_k(s) = \sum_{\nu=1}^{i} S_{kk}(s)\, C_{k\nu}(s)\, [X_\nu(s) - L_{\nu\nu}(s)\, Y_\nu(s)] +$$

$$+ \sum_{\mu=i+1}^{n} S_{kk}(s)\, C'_{k\mu}(s)\, [\Xi_\mu(s) - L_{\mu\mu}(s)\, W_\mu(s)] + V_k(s). \quad (5.12)$$

Соотношения (5.11) и (5.12) позволяют перейти к более компактному виду блок-схемы рассматриваемой системы по сравнению с блок-схемой, приведенной на фиг. 41.



Фиг. 42

Это компактное представление, указанное на фиг. 42, основано на введении единой матрицы системы, в которой входами служат сигналы рассогласований, а выходами — регулируемые переменные. Матрица ESC на фиг. 42 представляет собой матрицу, в которой элемент, расположенный на пересечении ν-го столбца и j-й строки, равен $E_{jk}\, S_{kk}\, C_{k\nu}$. Подобным же образом в матрице ESC' элемент, расположенный на пересечении j-й строки и μ-го столбца, равен $E_{jk}\, S_{kk}\, C'_{k\mu}$. Точно так же элементы матрицы SC равны $S_{kk}C_{k\nu}$, а элементы матрицы SC' равны $S_{kk}C'_{k\mu}$. Внешние возмущения V_k вводятся через посредство еще одной матрицы, образованной главным образом из элементов матрицы E двигателя.

5.3. Условия автономности. Теперь мы уже в состоянии дать аналитическое выражение критерия автономности регулирования. Задача заключается в определении условий, которые следует наложить на элементы $C_{k\nu}(s)$ и $C'_{k\mu}(s)$ матриц регулятора так, чтобы каждая из задающих функций $X_j(s)$ и $\Xi_\mu(s)$ влияла только на соответствующую ей переменную $Y_j(s)$ и $W_\mu(s)$, $j=1, 2, \ldots, i$; $\mu = i+1, \ldots, n$ и не влияла ни на какие другие переменные.

Так, например, задающая функция $X_2(s)$ должна изменять только выходную координату $Y_2(s)$, а задающая функция $\Xi_{i+1}(s)$ — только координату $W_{i+1}(s)$. Таким образом, в математическом отношении задача сводится к диагонализации матрицы системы, выписанной на фиг. 42. Мы формулируем критерий автономности в виде условий, налагаемых на элементы именно матрицы регулятора, по той причине, что эта матрица характеризует как раз те звенья полной системы, которые легче всего изменять в процессе проектирования системы. Что касается характеристик двигателя, исполнительных приводов и измерителей, то они считаются наперед заданными и проектировщик регулятора располагать ими не может.

Обратимся прежде всего к рассмотрению какого-либо одного выхода $Y_g(s)$, где индекс g может принимать одно из значений $1, 2, \ldots, i$. В этом случае равенства (5.11) и (5.12) примут вид

$$
\begin{aligned}
Y_j(s) = \sum_{k=1}^{n} \Bigg[& \sum_{\nu=1;\ \nu\neq g}^{i} E_{jk}\, S_{kk}\, C_{k\nu}(X_\nu - L_{\nu\nu} Y_\nu) + \\
& + \sum_{\mu=i+1}^{n} E_{jk}\, S_{kk}\, C'_{k\mu}(\Xi_\mu - L_{\mu\mu} W_\mu) + E_{jk} V_k \Bigg] + \\
& + \sum_{k=1}^{n} E_{jk}\, S_{kk}\, C_{kg}(X_g - L_{gg} Y_g)
\end{aligned}
$$

и

$$
\begin{aligned}
W_h(s) = & \sum_{\nu=1;\ \nu\neq g}^{i} S_{hk}\, C_{k\nu}(X_\nu - L_{\nu\nu} Y_\nu) + \\
& + \sum_{\mu=i+1}^{n} S_{kk} C'_{k\mu}(\Xi_\mu - L_{\mu\mu} W_\mu) + V_h + S_{kk} C_{kg}(X_g - L_{gg} Y_g).
\end{aligned}
$$

Для того чтобы задающая функция X_q не оказывала влияния на любую из координат Y_j или W_μ, за исключением Y_q, последние слагаемые правых частей последних двух соотношений должны равняться нулю при $j \neq g$ и при $k > i$. Таким образом, для любого g ($g = 1, 2, \ldots, i$) должны выполняться равенства

$$\sum_{k=1}^{n} E_{jk} S_{kk} C_{kg} = 0, \qquad j \neq g \tag{5.13}$$

и

$$C_{kg} = 0, \qquad k > i. \tag{5.14}$$

При выполнении условия (5.14) матрица регулятора значительно упрощается. Так, в примере системы, схематически показанной на фиг. 5.4, $i = 3$, $n = 5$. Тогда (5.14) требует, чтобы соблюдались равенства

$$C_{41} = C_{42} = C_{43} = C_{51} = C_{52} = C_{53} = 0.$$

С помощью условия (5.14) можно также упростить условие (5.13). В самом деле, условия (5.13) можно представить в виде

$$\sum_{k=1}^{i} E_{jk} S_{kk} C_{kg} = \sum_{k=1}^{i} \delta_{jg} E_{gk} S_{kk} C_{kg}, \tag{5.15}$$

где $g = 1, 2, \ldots, i$, а δ_{jg} — символ Кронекера, т. е.

$$\left.\begin{aligned} \delta_{jg} &= 0, \quad j \neq g, \\ \delta_{jg} &= 1, \quad j = g. \end{aligned}\right\} \tag{5.16}$$

При всяком фиксированном g уравнения (5.15), по существу, сводятся к системе $i-1$ линейных алгебраических уравнений относительно i неизвестных $S_{kk} C_{kg}$, $k = 1, 2, \ldots$ $\ldots, i$. Поэтому система определяет лишь отношения этих неизвестных, а не истинные значения самих неизвестных. Но это именно и есть то, что нам нужно, ибо мы не хотим фиксировать значение передаточной функции регулятора и тем самым потерять возможность варьирования тех или иных величин в процессе разработки регулятора.

Для нахождения этих отношений передаточных функций системы регулятора мы воспользуемся следующим свойством определителей: пусть $|E^*_{jl}|$ — алгебраическое до-

полнение элемента E_{jl} определителя $|E^*|$ квадратной матрицы E^*[1]). Тогда

$$\sum_{j=1}^{i} E_{jk} |E^*_{jl}| = 0, \qquad k \neq l$$

и (5.17)

$$\sum_{j=1}^{i} E_{jk} |E^*_{jl}| = |E^*|, \qquad k = l.$$

Умножив обе части уравнений (5.15) на $|E^*_{jl}|$ и произведя суммирование по индексу j, мы получим

$$\sum_{k=1}^{i} \sum_{j=1}^{i} |E^*_{jl}| \delta_{jg} E_{gk} S_{kk} C_{kg} = \sum_{k=1}^{i} \sum_{j=1}^{i} |E^*_{jl}| \cdot E_{jk} S_{kk} C_{kg}.$$

Следовательно, на основании соотношений (5.17) имеем

$$S_{ll} C_{lg} = \frac{|E^*_{gl}|}{|E^*|} \sum_{k=1}^{i} E_{gk} S_{kk} C_{kg}, \tag{5.18}$$

$$l = 1, 2, \ldots, i.$$

В частности,

$$S_{gg} C_{gg} = \frac{|E^*_{gg}|}{|E^*|} \sum_{k=1}^{i} E_{gk} S_{kk} C_{kg}.$$

Тогда, разделив последние два равенства одно на другое, придем к соотношению

$$\frac{S_{jj} C_{j\nu}}{S_{\nu\nu} C_{\nu\nu}} = \frac{|E^*_{\nu j}|}{|E^*_{\nu\nu}|}, \qquad j, \nu = 1, 2, \ldots, i, \tag{5.19}$$

посредством которого элементы матрицы SC, не лежащие на главной диагонали, выражаются через элементы, расположенные на этой диагонали.

Следовательно, соотношения (5.14) и (5.19) представляют собой необходимые условия автономности регулирования по переменным Y_g. Они были выведены Боксенбомом и Худом. Эти же авторы доказали также и доста-

[1]) Здесь символ $|E^*|$ обозначает не абсолютную величину E^*, а $\det E^*$. Рассматриваемое свойство определителей излагается в любом курсе высшей алгебры, например [31]. — *Прим. перев.*

точность упомянутых условий для автономности регулирования. Тем самым задача о разыскании матрицы C регулятора, при которой обеспечивается автономность регулирования, решена полностью.

Для решения задачи о разыскании другой части матрицы регулятора — матрицы C' — необходимо найти условия автономности регулирования по регулируемым переменным W_μ $(\mu = i+1, \ldots, n)$. С этой целью перепишем уравнения (5.11) и (5.12) в следующем виде:

$$Y_j(s) = \sum_{k=1}^{n} \Big[\sum_{\nu=1}^{i} E_{jk} S_{kk} C_{k\nu} (X_\nu - L_{\nu\nu} Y_\nu) + \sum_{\mu=i+1;\ \mu \neq r}^{n} E_{jk} S_{kk} C'_{k\mu} (\Xi_\mu - L_{\mu\mu} W_\mu) + E_{jk} V_k \Big] + \sum_{k=1}^{n} E_{jk} S_{kk} C'_{kr} (\Xi_r - L_{rr} W_r) \quad (5.20)$$

и

$$W_k(s) = \sum_{\nu=1}^{i} S_{kk} C_{k\nu} (X_\nu - L_{\nu\nu} Y_\nu) + \sum_{\substack{\mu=i+1 \\ \mu \neq r}}^{n} S_{kk} C'_{k\mu} (\Xi_\mu - L_{\mu\mu} W_\mu) + V_k + S_{kk} C'_{kr} (\Xi_r - L_{rr} W_r). \quad (5.21)$$

Здесь индекс r может принимать значения $i+1, \ldots, n$, а индекс j — значения $1, 2, \ldots, i$. В рассматриваемой задаче индекс k (5.21) пробегает значения $i+1, \ldots, n$, так как среди переменных $W_k(s)$ регулируемыми являются только те, для которых k пробегает именно эти значения. Как нетрудно видеть из рассмотрения уравнений (5.20) и (5.21), для того чтобы задающие функции Ξ_r влияли только на переменные W_r, последнее слагаемое в правой части каждого из этих уравнений должно обращаться в нуль, т. е.

$$\sum_{k=1}^{n} E_{jk} S_{kk} C'_{kr} = 0; \qquad j = 1, 2, \ldots, i \qquad (5.22)$$

$$C'_{kr} = 0 \text{ при } k, r = i+1, \ldots, n;\ k \neq r. \qquad (5.23)$$

Подобно тому, как это имело место выше, условие (5.23), аналогичное условию (5.14), позволяет немедленно упростить вид матрицы C'. Так, для системы, отвечающей схеме, приведенной на фиг. 41, $i=3$, $n=5$ и, в силу этого,

$$C'_{45} = C'_{54} = 0.$$

Условие (5.23) позволяет упростить также и уравнение (5.22), которое сводится к уравнению

$$\sum_{k=1}^{i} E_{jk} S_{kk} C'_{kr} = -E_{jr} S_{rr} C'_{rr}.$$

Умножая обе части последнего уравнения на алгебраические дополнения $|E^*_{jl}|$ и производя суммирование по j, получаем

$$\sum_{k=1}^{i} \sum_{j=1}^{i} |E^*_{jl}| E_{jk} S_{kk} C'_{kr} = -S_{rr} C'_{rr} \sum_{j=1}^{i} |E^*_{jl}| E_{jr}.$$

Пользуясь свойством определителей, выражаемым соотношениями (5.17), мы приходим к уравнению

$$|E^*| S_{ll} C'_{lr} = -S_{rr} C'_{rr} \sum_{j=1}^{i} |E^*_{jl}| E_{jr}.$$

Поэтому мы можем переписать последнее уравнение еще и иначе, заменяя в обозначениях индексов l на j и j на l:

$$\frac{S_{jj} C'_{jr}}{S_{rr} C'_{rr}} = -\frac{1}{|E^*|} \sum_{l=1}^{i} |E^*_{lj}| E_{lr}; \quad \begin{matrix} j = 1, \ldots, i \\ r = i+1, \ldots, n. \end{matrix} \tag{5.24}$$

Это соотношение определяет элементы матрицы SC', расположенные вне главной диагонали, через элементы, лежащие на этой диагонали. Соотношения (5.23) и (5.24) представляют собой необходимые и достаточные условия автономности регулирования по регулируемым переменным $W_\mu(s)$ $(\mu = i+1, \ldots, n)$.

Для обеспечения полной автономности по всем регулируемым переменным следует выполнить условия, выражаемые соотношениями (5.14), (5.19), (5.23) и (5.24). При этом элементы полной матрицы регулятора, не лежащие на главной диагонали, суть или нули, или величины, опре-

деляемые через посредство диагональных элементов. Таким образом, коль скоро характеристики двигателя заданы с помощью его матрицы, то диагональные элементы матрицы регулятора полностью определяют всю матрицу регулятора.

5.4. Уравнения, определяющие реакцию системы на выходе. При выполнении всех условий автономности уравнения (5.11) и (5.12) можно значительно упростить. Например, изменив порядок суммирования, получим

$$Y_j(s) = \sum_{\nu=1}^{l} [X_\nu(s) - L_{\nu\nu}(s) Y_\nu(s)] \sum_{k=1}^{n} E_{jk} S_{kk} C_{k\nu} +$$

$$+ \sum_{\mu=l+1}^{n} [\Xi_\mu(s) - L_{\mu\mu}(s) W_\mu(s)] \sum_{k=1}^{n} E_{jk} S_{kk} C'_{k\mu} + \sum_{k=1}^{n} E_{jk} V_k.$$

Но в силу условий автономности (5.13) и (5.14) регулирования вторая сумма (по k) в первом слагаемом правой части последнего равенства исчезает при всех ν и j, за исключением случая $\nu = j$. Второе слагаемое в правой части этого равенства исчезает, в силу условия автономности (5.22). В результате функция $Y_j(s)$ определяется с помощью соотношения

$$Y_j(s) = [X_j(s) - L_{jj}(s) Y_j(s)] \sum_{k=1}^{l} E_{jk} S_{kk} C_{kj} + \sum_{k=1}^{n} E_{jk} V_k.$$

Но, в силу (5.19), произведение $S_{kk} C_{kj}$ можно выразить посредством диагонального элемента $S_{jj} C_{jj}$. Выполнив это и использовав соотношения (5.17), мы придем к равенству

$$\sum_{k=1}^{l} E_{jk} S_{kk} C_{kj} = \frac{S_{jj} C_{jj}}{|E^*_{jj}|} \sum_{k=1}^{l} E_{jk} |E^*_{jk}| = S_{jj} C_{jj} \frac{|E^*|}{|E^*_{jj}|}.$$

Поэтому в окончательном виде получим

$$Y_j(s) = \frac{|E^*|}{|E^*_{jj}|} S_{jj} C_{jj} [X_j(s) - L_{jj}(s) Y_j(s)] + \sum_{k=1}^{n} E_{jk} V_k. \quad (5.25)$$

С помощью аналогичных вычислений уравнения (5.12) сводятся при использовании условий автономности регулиро-

вания к уравнениям

$$W_\mu(s) = S_{\mu\mu}C'_{\mu\mu}[\Xi_\mu(s) - L_{\mu\mu}(s)W_\mu(s)] + V_\mu(s), \quad (5.26)$$

$$\mu = i+1, \ldots, n.$$

Введя обозначения

$$R_{jj} = \frac{|E^*| S_{jj}C_{jj}}{E^* | S_{jj}C_{jj}L_{jj} + | E^*_{jj} |} \quad (5.27)$$

и

$$R'_{\mu\mu} = \frac{S_{\mu\mu}C'_{\mu\mu}}{S_{\mu\mu}C'_{\mu\mu}L_{\mu\mu} + 1}, \quad (5.28)$$

выпишем решения уравнений (5.25) и (5.26) относительно функций $Y_j(s)$ и $W_\mu(s)$ соответственно в виде

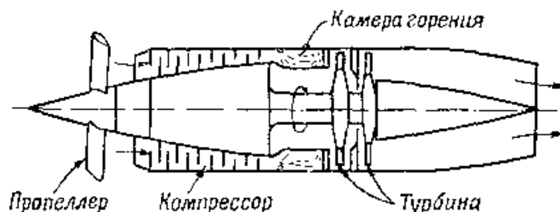
$$Y_j(s) = R_{jj}(s)X_j(s) - [R_{jj}(s)L_{jj}(s) - 1]\sum_{k=1}^{n} E_{jk}(s)V_k(s) \quad (5.29)$$

и

$$W_\mu(s) = R'_{\mu\mu}(s)\Xi_\mu(s) - [R'_{\mu\mu}(s)L_{\mu\mu}(s) - 1]V_\mu(s). \quad (5.30)$$

Равенства (5.29) и (5.30) служат для вычисления изображений регулируемых переменных по изображениям задающих воздействий и возмущений. По своему строению эти равенства совершенно аналогичны равенству (5.3) для одномерной системы автоматического регулирования. Функция $R_{jj}(s)$ служит полной передаточной функцией, связывающей выход $Y_j(s)$ со входом X_js. Функция $R'_{\mu\mu}(s)$ служит полной передаточной функцией, связывающей выход $W_\mu(s)$ со входом $\Xi_\mu(s)$. Эти полные передаточные функции определяются соотношениями (5.27) и (5.28), с помощью которых они выражаются через характеристики двигателя, исполнительных приводов, измерителей и регулятора. Разработка системы заключается в нахождении для каждого j и μ подходящей передаточной функции $C_{jj}(s)$ или $C'_{\mu\mu}(s)$, обеспечивающей удовлетворительное качество регулирования; эти функции находятся с помощью методов, изложенных в гл. IV. После этого с помощью соотношений (5.14), (5.19), (5.23) и (5.24) подсчитываются недиагональные элементы матрицы регулятора. Выполнив все эти операции, мы осуществим автономное регулирование сложной системы со многими регулируемыми параметрами, обладающее хорошим качеством.

5.5. Регулирование турбопропеллерного двигателя. В качестве простого примера применения общей теории автономного регулирования рассмотрим задачу о регулировании турбопропеллерного двигателя (фиг. 43). В процессе регулирования такого двигателя используются следующие переменные: угловая скорость турбопропеллерного агрегата, температура газов на входе турбины, угол установки лопастей пропеллера и количество топлива, подаваемое в единицу



Фиг. 43

времени в камеры сгорания. Систему регулирования двигателя следует спроектировать так, чтобы она обеспечивала работу двигателя в различных эксплуатационных установившихся режимах. Для каждого установившегося режима нужно исследовать качество переходных процессов в окрестности этого рабочего режима. Обозначим через $W_1(s)$ изображение по Лапласу отклонения угла установки лопастей пропеллера от его установившегося значения в данном рабочем режиме, а $W_2(s)$ — изображение по Лапласу аналогичного отклонения количества топлива, сжигаемого в единицу времени. Так как нас интересует течение переходных процессов в малой окрестности рабочего режима, мы можем линеаризовать соотношение между избытком создаваемого турбиной вращающего момента над моментом сопротивления, приложенным к компрессору и пропеллеру, углом установки лопаток и количеством топлива, сжигаемым в единицу времени. Поэтому упомянутый избыток крутящего момента будет выражаться линейной формой от $W_1(s)$ и $W_2(s)$. Обозначив еще через $Y_1(s)$ изображение по Лапласу отклонения угловой скорости от ее установившегося рабочего значения, выразим изображение избытка вращающего момента

с помощью произведения $(1+\tau s) Y_1(s)$, где τ — постоянная времени, характеризующая инерцию вращающихся масс двигателя [см. (4.1)]. Значение τ зависит от выбора рабочего режима двигателя[1]. Таким образом,

$$(1+\tau s) Y_1(s) = -aW_1(s) + bW_2(s), \qquad (5.31)$$

где a и b — вещественные положительные постоянные, подсчитываемые на основании протекания характеристик двигателя в малой окрестности рабочего режима. Коэффициенты a и b имеют определенный физический смысл: если количество топлива, сжигаемого в единицу времени, поддерживается неизменным, то $W_2(s) \equiv 0$; тогда из уравнения (5.31) следует, что $a = -Y_1(0)/W_1(0)$. Но значение $s=0$ отвечает установившемуся состоянию и, следовательно, величина a равна отношению уменьшения установившегося значения угловой скорости к увеличению угла установки лопастей пропеллера при неизменном режиме горения. Если построить семейство кривых, характеризующих изменение установившейся угловой скорости в зависимости от угла установки лопастей пропеллера при различных количествах топлива, сжигаемого в единицу времени[2], то a окажется равным угловому коэффициенту соответствующей кривой в выбранной рабочей точке.

Подобным же образом b равно угловому коэффициенту кривой, характеризующей изменение установившегося значения угловой скорости в зависимости от количества топлива, сжигаемого в единицу времени, при неизменном угле установки лопастей пропеллера опять-таки в рабочей точке[3]. Таким образом, постоянные a и b определяются статическими характеристиками двигателя.

В компрессорах осевого типа расход воздуха через компрессор при фиксированных условиях в области входного (заборного) отверстия остается почти постоянным при неизменной угловой скорости компрессора. Поэтому при данных условиях на входе в компрессор[1]) отношение

---

[1]) Ибо от числа оборотов двигателя зависит коэффициент c демпфирования в уравнении (4.1). — *Прим. перев.*

[2]) Такие кривые называются обычно статическими. — *Прим. перев.*

[3]) См. также [33, 41]. — *Прим. перев.*

количества тепла, сообщаемого массе газа, к этой массе газа является функцией от угловой скорости и количества топлива, подаваемого в камеры горения в единицу времени. Поэтому угловая скорость и режим горения определяют температуру на входе в турбину. Обозначим через $Y_2(s)$ изображение по Лапласу отклонения температуры на входе в турбину от ее рабочего значения. Тогда можно составить уравнение, связывающее изображения $Y_2(s)$, $W_1(s)$ и $W_2(s)$ и аналогичное уравнению (5.31) с той разницей, что постоянная времени, характеризующая время достижения термического равновесия, практически равна нулю, благодаря чему соотношение между этими изображениями оказывается более простым:

$$Y_2(s) = cW_2(s) - eY_1(s). \qquad (5.32)$$

Здесь c и e — также вещественные положительные постоянные. Постоянная c равна угловому коэффициенту касательной к статической характеристике температуры на входе в турбину по количеству топлива, сжигаемого в единицу времени, при постоянной угловой скорости ротора двигателя. Постоянная e равна угловому коэффициенту касательной к статической характеристике температуры на входе в турбину по угловой скорости ротора двигателя при неизменном расходе топлива в камерах горения; эти касательные проводятся к соответствующим кривым в рабочей точке.

Разрешив уравнения (5.31) и (5.32) относительно $Y_1(s)$ и $Y_2(s)$, придем к соотношениям

$$\left.\begin{aligned} Y_1(s) &= \frac{-a}{1+\tau s} W_1(s) + \frac{b}{1+\tau s} W_2(s), \\ Y_2(s) &= \frac{ae}{1+\tau s} W_1(s) + \frac{(c-be)+c\tau s}{1+\tau s} W_2(s), \end{aligned}\right\} \qquad (5.33)$$

определяющим матрицу E двигателя, рассматривавшуюся выше (см. стр. 94). Интересно отметить, что матрица E содержит только одну постоянную времени — постоянную τ.

---

[1]) Условия на входе в компрессор, т. е. на выходе диффузора, определяются аэродинамической скоростью, плотностью воздуха (высотой полета) и его температурой, числами Маха и Рейнольдса, угловой скоростью самолета и т. п. — *Прим. перев.*

Только эта постоянная времени характеризует свойства собственно двигателя. Полная же система автоматического регулирования обладает, конечно, и другими постоянными времени. Но эти постоянные входят через посредство характеристик регулятора, исполнительных приводов, измерителей и не содержатся в матрице двигателя.

Начнем с изучения задачи о регулировании угловой скорости ротора двигателя и расхода топлива в камерах горения. Таким образом, изображения регулируемых переменных суть $Y_1(s)$ и $W_2(s)$. Поэтому нам необходимо лишь первое из соотношений (5.33), причем $i=1$ и $n=2$. Отсюда следует, что матрица двигателя состоит только из двух элементов:

$$E_{11}=\frac{-a}{1+\tau s}, \quad E_{12}=\frac{b}{1+\tau s} \tag{5.34}$$

и

$$|E^*|=|E_{11}^*|E_{11}=E_{11}; \quad |E_{11}^*|=1. \tag{5.35}$$

Регулятор системы описывается соотношениями

$$\begin{aligned} U_1(s)=C_{11}(s)[X_1(s)-L_{11}(s)Y_1(s)]+ \\ +C'_{12}(s)[\Xi_2(s)-L_{22}(s)W_2(s)], \\ U_2(s)=C_{21}(s)[X_1(s)-L_{11}(s)Y_1(s)]+ \\ +C'_{22}(s)[\Xi_2(s)-L_{22}(s)W_2(s)]. \end{aligned} \tag{5.36}$$

Условия автономности требуют выполнения равенства

$$C_{21}(s)=0 \tag{5.37}$$

и, в силу (5.35), равенств

$$\frac{S_{11}(s)C'_{12}(s)}{S_{22}(s)C'_{22}(s)}=-\frac{|E_{11}^*|E_{12}}{|E^*|}=-\frac{E_{12}}{E_{11}}=\frac{b}{a}. \tag{5.38}$$

Так как величина $-a$ равна частной производной от угловой скорости ротора двигателя по углу установки лопастей пропеллера, а величина b равна частной производной от угловой скорости по расходу топлива в камерах горения, отношение b/a определяет быстроту изменения угла установки лопаток пропеллера по изменению расхода топлива в камерах горения при постоянстве угловой скорости ротора двигателя. Конечно, отношение b/a зависит

от параметров полета самолета, снабженного турбопропеллерным двигателем. Так, например, это отношение увеличивается с ростом высоты полета. Поэтому в правильно спроектированном регуляторе должны быть предусмотрены средства компенсации влияний, вызванных изменением режима полета и рабочего режима двигателя.

Таким образом, передаточная функция $R_{11}(s)$, характеризующая выходную скорость ротора двигателя, определяется выражением

$$R_{11}(s) = \frac{aS_{11}(s)\,C_{11}(s)}{aS_{11}(s)\,C_{11}(s)\,L_{11}(s) + (1 + \tau s)}, \qquad (5.39)$$

а передаточная функция $R'_{22}(s)$, характеризующая выходной расход топлива в камерах горения, определяется выражением

$$R'_{22}(s) = \frac{S_{22}(s)\,C'_{22}(s)}{S_{22}(s)\,C'_{22}(s)\,L_{22}(s) + 1}. \qquad (5.39')$$

Правые части последних соотношений определяют передаточные функции угловой скорости ротора двигателя и расхода топлива в камерах горения при автономном регулировании. Задача заключается в синтезе таких передаточных функций $C_{11}(s)$ и $C'_{22}(s)$ регулятора, при которых качество регулирования удовлетворяло бы поставленным требованиям во всей области ожидаемых рабочих режимов.

Обратимся теперь к рассмотрению другого возможного способа регулирования турбопропеллерного двигателя. Этот способ основан на регулировании угловой скорости ротора и температуры газов на входе в турбину. В этом случае изображения регулируемых переменных суть $Y_1(s)$ и $Y_2(s)$. Следовательно, здесь мы нуждаемся в обоих соотношениях (5.33) и $i = n = 2$. Условия автономности принимают вид:

$$\left.\begin{aligned} \frac{S_{22}(s)\,C_{21}(s)}{S_{11}(s)\,C_{11}(s)} &= -\frac{ae}{(c-be)+c\tau s}, \\ \frac{S_{11}(s)\,C_{12}(s)}{S_{22}(s)\,C_{22}(s)} &= \frac{b}{a}. \end{aligned}\right\} \qquad (5.40)$$

Передаточная функция $R_{11}(s)$ по угловой скорости ротора

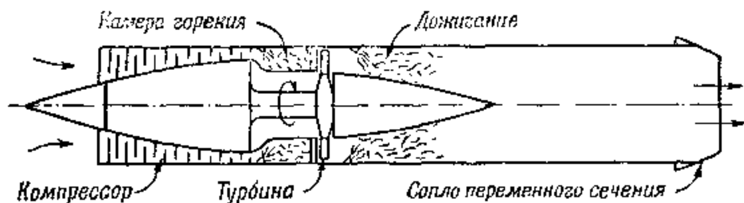
определяется, следовательно, выражением

$$R_{11}(s) = \frac{S_{11}(s)\,C_{11}(s)}{S_{11}(s)\,C_{11}(s)\,L_{11}(s) - \dfrac{(c-be)+c\tau s}{ac}}, \tag{5.41}$$

а передаточная функция $R_{22}(s)$ по температуре на входе в турбину — выражением

$$R_{22}(s) = \frac{S_{22}(s)\,C_{22}(s)}{S_{22}(s)\,C_{22}(s)\,L_{22}(s) + \dfrac{1}{c}}.$$

5.6. Турбореактивный двигатель с дожиганием. Рассмотрим теперь задачу регулирования турбореактивного двигателя с дожиганием, упомянутую в начале этой главы. Схема двигателя изображена на фиг. 44. Мы опять будем



Фиг. 44

рассматривать задачу регулирования для неустановившихся процессов в малой окрестности выбранной рабочей точки. Поэтому можно применить линеаризацию соотношений, связывающих различные переменные.

Пусть, как и прежде, $Y_1(s)$ означает изображение (по Лапласу) отклонения угловой скорости ротора двигателя от ее рассматриваемого установившегося значения, $W_1(s)$ — изображение отклонения проходного сечения для истечения газов, $W_2(s)$ — изображение отклонения расхода топлива в камерах горения, $W_3(s)$ — изображение отклонения расхода топлива при дожигании; все эти отклонения отсчитываются от рассматриваемого установившегося режима. С помощью этих обозначений составим уравнение, аналогичное уравнению (5.31) для турбопропеллера:

$$(1+\tau s)Y_1(s) = a_1W_1(s) + a_2W_2(s) + a_3W_3(s), \tag{5.42}$$

где a_1, a_2 и a_3 — вещественные постоянные. Как и в случае турбопропеллерного двигателя, эти постоянные представляют собой угловые коэффициенты касательных к статическим характеристикам двигателя. Так, постоянная a_1 равна быстроте изменения угловой скорости ротора двигателя при изменении открытия пропускного хвостового отверстия для выхода газов при неизменном расходе топлива в камерах горения и на дожигании. Постоянная a_2 равна быстроте изменения угловой скорости ротора двигателя при изменении расхода топлива в камерах горения. Постоянная a_3 равна быстроте изменения угловой скорости ротора двигателя при изменении расхода топлива при дожигании. Величина τ, фигурирующая в уравнении (5.42), представляет собой, как и выше, единственную постоянную времени двигателя; τ характеризует инерцию вращающихся масс двигателя. Рассмотренное линеаризированное соотношение между угловой скоростью двигателя и другими входными переменными двигателя было получено М. С. Федером и Р. Худом [1] [2].

Если в двигателе используется компрессор осевого типа, то мы можем вновь применить уравнение (5.32) предыдущего раздела. Обозначив через $Y_2(s)$ изображение температуры газов на входе в турбину, напишем

$$Y_2(s) = -eY_1(s) + cW_2(s).$$

Разрешив последнее уравнение и уравнение (5.42) относительно $Y_1(s)$ и $Y_2(s)$, получим

$$\left.\begin{aligned} Y_1(s) &= \frac{a_1}{1+\tau s} W_1(s) + \frac{a_2}{1+\tau s} W_2(s) + \frac{a_3}{1+\tau s} W_3(s), \\ Y_2(s) &= -\frac{a_1 e}{1+\tau s} W_1(s) + \frac{(c-a_2 e)+c\tau s}{1+\tau s} W_2(s) - \frac{a_3 e}{1+\tau s} W_3(s). \end{aligned}\right\} \tag{5.43}$$

Таким образом, элементы матрицы двигателя определяются

[1] Feder M. S., Hood R., NACA TN 2183 (1950).

[2] Угловая скорость ротора двигателя является одной из его выходных переменных. — *Прим. перев.*

соотношениями

$$\left.\begin{aligned} E_{11}=\frac{a_1}{1+\tau s}; \quad E_{12}=\frac{a_2}{1+\tau s}; \quad E_{13}=\frac{a_3}{1+\tau s}, \\ E_{21}=-\frac{a_1 e}{1+\tau s}; \quad E_{22}=\frac{(c-a_2 e)+c\tau s}{1+\tau s}; \quad E_{23}=-\frac{a_3 e}{1+\tau s}. \end{aligned}\right\} \tag{5.44}$$

Рассмотрим задачу об автоматическом регулировании угловой скорости ротора двигателя, температуры газов на входе в турбину и расхода топлива на дожигание. Таким образом, изображения регулируемых переменных суть $Y_1(s)$, $Y_2(s)$ и $W_3(s)$. Уравнения процесса регулирования имеют вид

$$\left.\begin{aligned} U_1(s)=C_{11}(s)[X_1(s)-L_{11}(s)Y_1(s)]+ \\ +C_{12}(s)[X_2(s)-L_{22}(s)Y_2(s)]+C'_{13}(s)[\Xi_3(s)-L_{33}(s)W_3(s)], \\ U_2(s)=C_{21}(s)[X_1(s)-L_{11}(s)Y_1(s)]+ \\ +C_{22}(s)[X_2(s)-L_{22}(s)Y_2(s)]+C'_{23}(s)[\Xi_3(s)-L_{33}(s)W_3(s)], \\ U_3(s)=C_{31}(s)[X_1(s)-L_{11}(s)Y_1(s)]+ \\ +C_{32}(s)[X_2(s)-L_{22}(s)Y_2(s)]+C'_{33}(s)[\Xi_3(s)-L_{33}(s)W_3(s)], \end{aligned}\right\} \tag{5.45}$$

где $X_1(s)$, $X_2(s)$ и $\Xi_3(s)$ обозначают изображения заданных значений угловой скорости ротора двигателя, температуры газов на входе в турбину и расхода топлива на дожигание соответственно.

Применяя условие (5.14) автономности регулирования, мы сразу же видим, что

$$C_{31}(s)=C_{32}(s)=0. \tag{5.46}$$

Условия (5.19) сводятся к соотношениям:

$$\left.\begin{aligned} \frac{S_{11}(s)C_{12}(s)}{S_{22}(s)C_{22}(s)}=-\frac{a_2}{a_1}, \\ \frac{S_{22}(s)C_{21}(s)}{S_{11}(s)C_{11}(s)}=\frac{a_1 e}{(c-a_2 e)+c\tau s}. \end{aligned}\right\} \tag{5.47}$$

Условие автономности (5.24) приводит к соотношениям:

$$\begin{aligned} \frac{S_{11}(s)C'_{13}(s)}{S_{33}(s)C'_{33}(s)}=-\frac{a_3}{a_1}, \\ C'_{23}(s)=0. \end{aligned} \tag{5.48}$$

Отношения $-a_2/a_1$ и $-a_3/a_1$ в равенствах (5.47) и (5.48) допускают простое физическое истолкование: $-a_2/a_1$ определяет быстроту изменения выходного сечения двигателя с изменением расхода топлива в камерах горения при постоянстве угловой скорости ротора двигателя и расхода топлива на дожигание, $-a_3/a_1$ определяет быстроту изменения выходного сечения двигателя с изменением расхода топлива на дожигание при постоянстве угловой скорости ротора двигателя и расхода топлива в камерах горения.

При выполнении соотношений (5.46) – (5.48) процесс автоматического регулирования будет автономным, и передаточная функция $R_{11}(s)$ по угловой скорости двигателя выразится в форме

$$R_{11}(s) = \frac{S_{11}(s)\,C_{11}(s)}{S_{11}(s)\,C_{11}(s)\,L_{11}(s) + \frac{(c-a_2e)+c\tau s}{a_1c}}. \qquad (5.49)$$

Передаточная функция $R_{22}(s)$ по температуре на входе в турбину имеет вид

$$R_{22}(s) = \frac{S_{22}(s)\,C_{22}(s)}{S_{22}(s)\,C_{22}(s)\,L_{22}(s) + \frac{1}{c}}, \qquad (5.50)$$

а передаточная функция $R'_{33}(s)$ по расходу топлива на дожигание равна

$$R'_{33}(s) = \frac{S_{33}(s)\,C'_{33}(s)}{S_{33}(s)\,C'_{33}(s)\,L_{33}(s)+1}. \qquad (5.51)$$

Последние соотношения служат отправной точкой при синтезе передаточных функций $C_{11}(s)$, $C_{22}(s)$, $C'_{33}(s)$ и, следовательно, передаточных функций $C_{12}(s)$, $C_{21}(s)$ и $C'_{13}(s)$.

Глава VI

СЛЕДЯЩИЕ СИСТЕМЫ НА ПЕРЕМЕННОМ ТОКЕ И СЛЕДЯЩИЕ СИСТЕМЫ КОЛЕБАТЕЛЬНОГО УПРАВЛЕНИЯ

Настоящая глава и две следующих посвящены вопросам обобщения понятий и методов, развитых в гл. II и III применительно к простым следящим системам, на линейные системы более сложного вида, которые тем не менее можно *приближенно* исследовать с помощью этих же методов. Возможность такого обобщения служит наглядной иллюстрацией силы основных принципов, на которых основан синтез следящих систем. При изложении материала этой и следующих глав мы будем придерживаться книги Маккола[1]).

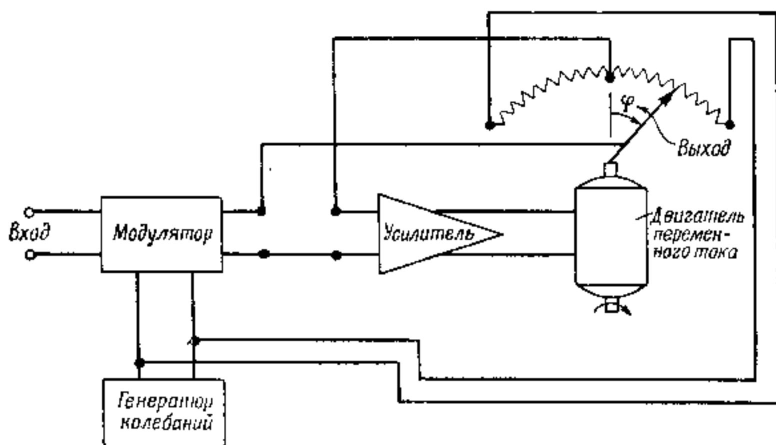
6.1. Системы на переменном токе. В предшествующих главах при рассмотрении систем автоматики, содержащих электрический двигатель, мы каждый раз неявно подразумевали двигатель постоянного тока. Однако в практических условиях подчас весьма желательно применять двигатели переменного тока. Очевидно, что введение в систему таких двигателей требует пересмотра некоторых частей нашего предшествующего рассуждения.

Рассмотрим следящую систему, схематически изображенную на фиг. 45. Назначение системы состоит в осуществлении поворота ротора двигателя на угол φ в соответствии со входным сигналом. Выходной угол φ измеряется с помощью потенциометра. Напряжение, снимаемое с потенциометра, служит сигналом обратной связи. В этой системе все электрические токи и напряжения в усилителе, двигателе и потенциометре суть модулированные синусоиды, т. е. синусоидальные функции постоянной частоты, равной ω_0, но с изменяющейся во времени

---

[1]) Маккол Л. А., Основы теории сервомеханизмов, М., 1947.

амплитудой. Опорное переменное напряжение вырабатывается генератором синусоидальных электрических колебаний. При выполнении определенного условия, к выяснению которого мы сейчас приступим, к такой системе можно применить значительную часть ранее изложенной теории.



Фиг. 45

Предположим на мгновение, что вся общая теория установившихся режимов в линейных системах с постоянными параметрами рассматривается в условиях воздействия на эти системы сигналов, представляющих собой модулированные синусоиды. При этом, говоря об «установившихся режимах», мы здесь имеем в виду то обстоятельство, что в качестве модулирующих сигналов принимаются чисто синусоидальные функции времени. Пусть несущая частота определяется (до модуляции) функцией $\cos \omega_0 t$; фазовый угол мы здесь опускаем без потери общности. Поскольку закон изменения несущей выражен в вещественной форме, модулирующий сигнал целесообразно записать в комплексной форме: $e^{i\omega t}$. После этого выражение для модулированной несущей принимает вид

$$x(t) = e^{i\omega t} \cos \omega_0 t = \frac{1}{2} [e^{i(\omega+\omega_0)t} + e^{i(\omega-\omega_0)t}], \qquad (6.1)$$

и, в силу (2.16), установившаяся реакция $y_{\text{уст.}}(t)$ системы, обладающей передаточной функцией $F(s)$, определяется выражением

$$y_{\text{уст.}}(t) = \frac{1}{2}\left[F(i\omega + i\omega_0)\, e^{i\omega_0 t} + F(i\omega - i\omega_0)\, e^{-i\omega_0 t}\right] e^{i\omega t}. \qquad (6.2)$$

В реальных системах функция $F(s)$ обычно представляет собой отношение многочленов по степеням s, имеющих вещественные коэффициенты. Тогда, как уже отмечалось, имеет место соотношение (3.17)

$$F(-i\omega) = \overline{F(i\omega)}, \qquad (6.3)$$

где черта над обозначением комплексной величины попрежнему означает комплексно-сопряженную величину. Поэтому правую часть равенства (6.2) можно переписать следующим образом:

$$\frac{1}{2}\left[F^*(i\omega)\, e^{i\omega_0 t} + \overline{F^*(-i\omega)}\, e^{-i\omega_0 t}\right] e^{i\omega t}, \qquad (6.4)$$

где

$$F^*(i\omega) = F(i\omega + i\omega_0). \qquad (6.5)$$

Будем считать систему такой, что справедливо соотношение

$$F(i\omega_0 + i\omega) = \overline{F(i\omega_0 - i\omega)}. \qquad (6.6)$$

В этом случае функцию (6.4) можно записать в виде

$$F^*(i\omega)\, e^{i\omega t} \cos \omega_0 t.$$

Этот результат показывает, что коль скоро выполнено условие (6.5), амплитуда реакции системы на входное воздействие, определяемое модулированной несущей (6.1), совпадает с реакцией системы, обладающей частотной характеристикой $F^*(i\omega)$ при входной частоте ω. Это утверждение допускает непосредственное обобщение на случай подачи на вход системы функций более общего вида, что отвечает принципу наложения, справедливому для линейных систем. Если условие (6.6) выполняется хотя бы приближенно в диапазоне частот ω, включающем наиболее важные части спектра Фурье модулирующего входного сигнала $x(t)$, то амплитуда соответствующего моду-

лирующего выходного сигнала равна (по крайней мере приблизительно) реакции на этот же вход $x(t)$ в системе, обладающей частотной характеристикой $F^*(i\omega)$. В гл. IV мы показали, что качество следящей системы полностью определяется ее частотной характеристикой. Приближенное же значение частотной характеристики равно теперь $F^*(i\omega)$. Тем самым открывается возможность применения к системам на переменном токе всех методов определения качества следящих систем, развитых в гл IV. Единственное различие заключается в использовании функции $F^*(i\omega)$ вместо функции $F(i\omega)$.

6.2. Перенос передаточной функции. Если оставить в стороне некоторые тривиальные системы, как, например, простые сопротивления, то из свойства (6.3) следует, что

$$F(i\omega_0 + i\omega) = \overline{F(-i\omega_0 - i\omega)}.$$

Это соотношение отличается от условия (6.6). Поэтому условие (6.6) нельзя в точности физически осуществить при всех вещественных значениях ω. Но, слегка изменив способ подхода к задаче, мы можем сказать, что частотные характеристики $F(i\omega)$ и $F^*(i\omega)$ двух *реальных* физических систем не могут в точности удовлетворять соотношению (6.5) при всех вещественных значениях ω. Тем не менее вполне возможно и это обычно делают, — чтобы частотные характеристики $F^*(i\omega)$ и $F(i\omega)$ двух реальных физических систем удовлетворяли соотношениям (6.5) приближенно на некотором диапазоне частот ω, достаточно широком в смысле включения наиболее важных частей спектра Фурье интересующих нас конкретных видов входных сигналов. Это можно вкратце показать следующим образом.

Рассмотрим импеданс Z цепи, состоящей из последовательного соединения индуктивности L и емкости C и питаемой током частоты ω':

$$Z = Li\omega' + \frac{1}{Ci\omega'} = Li\omega' \left(1 - \frac{1}{LC\omega'^2}\right)$$

Выбрав L и C так, чтобы имело место соотношение

$$\omega_0^2 = \frac{1}{LC}, \tag{6.7}$$

получим

$$Z = Li\omega' \left(1 - \frac{\omega_0^2}{\omega'^2}\right) = Li(\omega' - \omega_0)\left(1 + \frac{\omega_0}{\omega'}\right).$$

При малых значениях разности $\omega' - \omega_0 = \omega$, т. е. при значениях $\omega + \omega_0$, близких к ω_0,

$$Z \approx 2Li(\omega' - \omega_0) = 2Li\omega.$$

Это означает, что при частоте $\omega' = \omega_0 + \omega$ импеданс последовательно соединенных L и C, величины которых удовлетворяют равенству (6.7), приблизительно равен импедансу индуктивности $2L$ при частоте ω.

Подобным же образом рассмотрим импеданс Z цепи, образованной параллельным соединением индуктивности L и емкости C при частоте ω':

$$\frac{1}{Z} = \frac{1}{Li\omega'} + Ci\omega' \approx 2Ci(\omega' - \omega_0) = 2Ci\omega$$

при выполнении условия (6.7). Это означает, что при частоте $\omega' = \omega + \omega_0$ импеданс параллельно соединенных L и C, величины которых удовлетворяют равенству (6.7), приблизительно равен импедансу емкости $2C$ при частоте ω.

Что касается импеданса чистого активного сопротивления, то он, конечно, не зависит от частоты и имеет одно и то же значение как при частоте $\omega + \omega_0$, так и при частоте ω. Таким образом, отправляясь от реальной физической системы, обладающей передаточной функцией $F^*(s)$, мы можем построить реальную физическую систему, обладающую передаточной функцией $F(s)$, для которой условие (6.6) приближенно выполняется при малых значениях ω с помощью следующего приема: каждая индуктивность L заменяется последовательным соединением индуктивности $L/2$ и емкости $C = 2/(L\omega_0^2)$, а каждая емкость C — параллельным соединением емкости $C/2$ и индуктивности $L = 2/(C\omega_0^2)$. Этот прием перехода от функции $F^*(s)$ к функции $F(s)$ называется переносом передаточной функции по шкале частот на величину ω_0.

Пусть ω_0 — частота несущей, вырабатываемой генератором электрических колебаний. Очевидно, что все токи и напряжения в рассматриваемой системе существуют в форме соответствующих модулированных колебаний несущей, изменяющейся по закону $\cos \omega_0 t$. Из этого обстоятельства сразу же вытекает, что разработка усилителя для системы на переменном токе сводится к разработке с помощью методов гл. IV подходящего усилителя для системы на постоянном токе с последующим переносом его характеристики вверх по шкале частот на величину ω_0 по способу, разобранному выше.

Как уже отмечалось, все эти рассуждения связаны со значительным числом различных приближений и допущений. Для разработки полностью законченной теории следящих систем на переменном токе необходимо оценить влияние всех этих приближений и допущений. Мы не будем здесь заниматься таким исследованием ввиду его сложности и трудоемкости, а также по той причине, что, с точки зрения задач проектирования следящих систем, это исследование не представляется неотложным.

6.3. Следящие системы колебательного управления. Рассмотрим теперь другой класс систем, которые будем называть следящими системами колебательного управления. Такие системы схожи со следящими системами с двигателями на переменном токе в том отношении, что в обоих случаях периодическое колебание модулируется определенными сигналами. Однако модуляция, применяемая в случае следящих систем колебательного управления, не сводится к обыкновенной амплитудной модуляции. Чтобы идея следящей системы колебательного управления была понятной, сделаем краткие предварительные замечания.

Один весьма примитивный, но довольно распространенный тип следящей системы можно описать следующим образом: предположим, что к двигателю переменного тока подключается цепь, содержащая реле. Последнее подает напряжение на зажимы двигателя только тогда, когда абсолютное значение входной переменной $x(t)$ превышает известный порог, и при этом таким образом, что при превышении величиной $|x|$ этого порога двигатель

оказывается под полным напряжением E источника электродвижущей силы; полярность же включения этого источника должна быть такова, чтобы соответствующее перемещение ротора двигателя приводило к уменьшению абсолютной величины рассогласования. Этот пример соответствует следящей системе, действующей по так называемому принципу «да – нет».

Большое преимущество таких следящих систем состоит в том, что, будучи сравнительно простыми, они могут применяться для управления большими мощностями, чего часто трудно достигнуть с помощью следящих систем других типов. С другой стороны, следящие системы, действующие по принципу «да — нет», являются существенно нелинейными системами; как будет показано в гл. X, такие следящие системы действуют в общем хуже, чем системы, рассмотренные ранее. Короче говоря, следящая система колебательного управления представляет собой такое видоизменение следящей системы, действующей по принципу «да – нет», которое обеспечивает сохранение преимущества, связанного с линейностью характеристик, без потерь преимущества, связанного со способностью управления большими мощностями.

Перед тем как приступить к изучению самих следящих систем колебательного управления, приведем один общий теоретический результат, на котором построена теория всех таких систем. Рассмотрим устройство, обладающее следующими свойствами: в зависимости от того, является ли входной сигнал $x(t)$ положительным или отрицательным, выходной сигнал $y(t)$ равен $+A$ или $-A$, где A – постоянная величина. Такое устройство можно считать идеальным реле, т. е. системой с нулевым пороговым значением[1]). Пусть на вход реле подается сигнал, определяемый выражением

$$x(t) = E_0 \sin \omega_0 t + kE_0 \sin \omega t, \tag{6.8}$$

где E_0, k, ω_0 и ω постоянные. В следящей системе колебательного управления слагаемое $E_0 \sin \omega_0 t$ характеризует установившиеся колебания системы, а слагаемое

---

[1]) Или, иначе, с нулевой зоной нечувствительности. — *Прим. ред.*

$kE_0 \sin \omega t$ характеризует приложенный, или модулирующий сигнал. Перейдем теперь к подсчету соответствующего выходного сигнала $y(t)$.

6.4. Частотная характеристика реле. Выходной сигнал релейного устройства, отвечающий входному сигналу (6.8), можно представить в виде

$$\sum_{m=0}^{\infty} \sum_{n=-\infty}^{\infty} a_{mn} \sin [(m\omega_0 + n\omega) t], \tag{6.9}$$

где коэффициенты a_{mn} не зависят от времени. При $m = 0$ внутреннее суммирование следует производить только по положительным значениям n. Для нашей задачи непосредственно интересны только коэффициенты a_{10} и a_{01}, ибо в следящей системе колебательного управления, работающей в нормальных условиях, остальные коэффициенты либо пренебрежимо малы, либо соответствуют колебаниям, подавляемым с помощью надлежащих фильтров.

При $k = 0$, т. е. когда выходной сигнал релейного устройства представляет собой чисто синусоидальную функцию частоты ω_0, выходная переменная этого реле изменяется по закону бесконечной прямоугольной волны периода $\omega_0/4\pi$, причем абсолютные величины высот положительных и отрицательных волн равны общей постоянной A. Как известно, амплитуда a_{01} первой гармоники в разложении такой волны в ряд Фурье равна

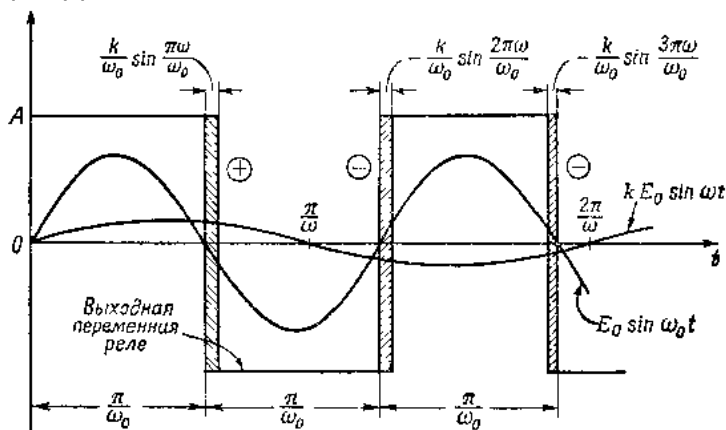
$$a_{01} = \frac{4A}{\pi}. \tag{6.10}$$

При $k \neq 0$ выходная переменная реле имеет вид, графически изображенный на фиг. 46. Разность между значениями выходных переменных реле, отвечающих случаям $k \neq 0$ и $k = 0$, графически изображается последовательностью прямоугольников высоты $2A$, заштрихованных на фиг. 46. При $|k| \ll 1$ отклонение моментов изменения знака выходной переменной реле от моментов, соответствующих равноотстоящим значениям времени $t_n = n\pi/\omega_0$, очень мало. Таким образом, прямоугольники, характеризующие это отклонение, очень узки, что и показано на фигуре. Ширину этих прямоугольников можно

определить с помощью отношения значения модулирующего сигнала при $t=t_n$ к угловому коэффициенту касательной, проведенной к кривой установившихся колебаний в той же точке $t=t_n$, т. е. искомая ширина равна

$$\left|\frac{kE_0 \sin \omega t_n}{E_0 \cos \omega_0 t_n}\right| = \frac{k}{\omega_0} |\sin \omega t_n|.$$

Рассмотренные прямоугольные площади складываются (+) с площадями соответствующих волн при немодулированном выходе или вычитаются (−) из них в зависимости от того, будет ли величина $\sin \omega t_n$ положительной



Фиг. 46

или отрицательной. Следовательно, площади прямоугольников можно измерить посредством величины

$$\frac{2Ak}{\omega_0} \sin \omega t_n.$$

Коэффициент a_{10} в выражении (6.9) является коэффициентом при первой гармонике в разложении в ряд Фурье последовательности этих узких прямоугольных волн. Так как за значение $\sin \omega t$ в области этих прямоугольников можно принять величины $\sin \omega t_n$, то, рассматривая N

таких поправочных прямоугольников, получаем

$$a_{10} \int_{0}^{N\pi/\omega_0} \sin^2 \omega t \, dt = 2A \frac{k}{\omega_0} \sum_{n=0}^{N} \sin^2 \omega t_n.$$

Но

$$\int_{0}^{N\pi/\omega_0} \sin^2 \omega t \, dt = \frac{1}{2} \int_{0}^{N\pi/\omega_0} (1 - \cos 2\omega t) \, dt =$$

$$= \frac{1}{2} \frac{N\pi}{\omega_0} - \frac{1}{4\omega} \sin \left(2N\pi \frac{\omega}{\omega_0} \right)$$

и

$$\sum_{n=0}^{N} \sin^2 \omega t_n = \sum_{n=1}^{N} \sin^2 \left(n\pi \frac{\omega}{\omega_0} \right) =$$

$$= \frac{1}{2} \sum_{n=1}^{N} \left[1 - \cos \left(2n\pi \frac{\omega}{\omega_0} \right) \right] = \frac{N}{2} - \frac{1}{2} \sum_{n=1}^{N} \cos \left(2n\pi \frac{\omega}{\omega_0} \right)$$

Последняя сумма остается ограниченной при неограниченном увеличении N. Поэтому, выбрав число N достаточно большим, будем иметь

$$a_{10} = 2 \frac{Ak}{\pi}. \tag{6.11}$$

Соотношения (6.10) и (6.11) определяют важные коэффициенты a_{01} и a_{10} при малых k. При произвольных значениях k эти коэффициенты вычислили Р. М. Кальб и У. Р. Беннетт[1]). При $0 < k < 1$

$$\left. \begin{aligned} a_{01} &= \frac{8A}{\pi^2} E(k), \\ a_{10} &= \frac{8A}{\pi^2 k} [E(k) - (1 - k^2) K(k)], \end{aligned} \right\} \tag{6.12}$$

где $K(k)$ и $E(k)$ означают полные эллиптические интегралы соответственно первого и второго рода. При малых k эллиптические интегралы можно разложить в ряды;

[1]) Kalb R. M., Bennett W. R., Bell System Tech. J., 14, 322—359 (1935).

тогда

$$\left.\begin{aligned} a_{01} &= \frac{4A}{\pi}\left(1 - \frac{k^2}{4} - \ldots\right), \\ a_{10} &= \frac{2Ak}{\pi}\left(1 + \frac{k^2}{8} + \ldots\right). \end{aligned}\right\} \tag{6.13}$$

Как показывают соотношения (6.13), наш простой метод подсчета a_{01} и a_{10} оказался верным в пределах принятой степени точности. Но они указывают также и на то, что простые равенства (6.10) и (6.11) в достаточной мере точны даже и при умеренно больших значениях k. Поэтому отношение составляющей выходной переменной, имеющей частоту ω, к составляющей с такой же частотой входной переменной, т. е. частотная характеристика $F_i(i\omega)$, приближенно определяется равенством

$$F_i(i\omega) = \frac{2A}{\pi E_0}. \tag{6.14}$$

Согласно (6.13) при малых k амплитуду составляющей выходного сигнала, имеющей частоту ω_0, можно приближенно считать постоянной, определяемой всецело свойствами реле. Кроме того, отношение составляющей частоты ω_0 выходной переменной к составляющей этой же частоты входной переменной равно $4A/(\pi E_0)$. Таким образом, усиление реле по составляющей с частотой ω_0 на 6 *дб* превышает усиление по составляющей с частотой ω.

Теперь становится очевидным, что предыдущие рассуждения можно обобщить на случай, когда вместо сигнала $kE_0 \sin\omega t$ мы имеем сигнал $x(t)$ произвольной формы, амплитуда которого гораздо меньше амплитуды E_0 установившегося колебания. Основной результат для этого случая можно сформулировать следующим образом: если пренебречь модуляционными членами более высокого порядка на том основании, что они ничтожно малы или подлежат, в конечном счете, подавлению посредством надлежащей фильтрации, то поведение реле, находящегося под воздействием входного сигнала

$$E_0 \sin\omega_0 t + x(t),$$

где $x(t)$ мало по сравнению с E_0, в основном подобно — поскольку речь идет о передаче входного сигнала — поведению линейной системы, обладающей частотной характеристикой (6.14).

6.5. Следящие системы колебательного управления, модулируемые колебаниями, снимаемыми с выхода этих систем. Теперь мы подготовлены для дальнейшего исследования следящих систем колебательного управления. Мы уже видели, что в части передачи сигналов единственное следствие введения в систему установившихся колебаний выражается в превращении релейного устройства в подлинно линейное звено, обладающее положительной вещественной частотной характеристикой. Отсюда следует вывод о возможности с самого начала рассматривать реле в качестве такого звена, без всякого упоминания в явной форме о наличии в системе колебания $E_0 \sin \omega_0 t$, и далее подходить к исследованию этой системы исключительно с помощью методов и понятий, изложенных в предыдущих главах. Эту поистине новую и превосходную методику предложил Ж. С. Лозье[1][2].

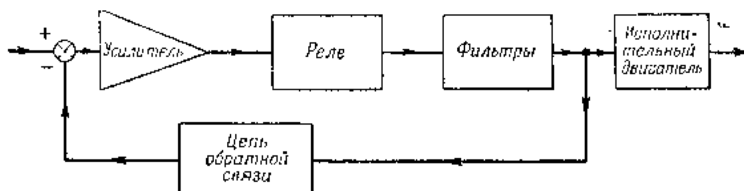
До сих пор мы предполагали в целях простоты, что системе присущи фильтрующие свойства, необходимые для подавления нежелательных модуляционных гармоник, поступающих в систему через посредство реле. Но понятно, что в практических условиях иногда требуется вводить в систему дополнительные фильтры для повышения ее фильтрующих свойств. Разумеется, что каковы бы ни были фильтры, имеющиеся в системе, они должны обладать способностью пропускать полезные сигналы. Это обстоятельство, рассматриваемое вместе с другими нашими заключениями, приводит нас к выводу о том, что частота ω_0 должна быть выше всех частот, лежащих в важных частях спектра Фурье передаваемых сигналов.

---

[1]) Математическое обоснование метода линеаризации нелинейных систем путем наложенных колебаний имеется в книге Н. Н. Боголюбова «О некоторых статистических методах в математической физике», АН УССР, 1945, стр. 7. — *Прим. перев.*

[2]) Эквивалентная линеаризация рассматривалась в работах А. А. Красовского (Автоматика и Телемеханика, 9, 1 (1948)) и Г. С. Поспелова (Труды ВВА им. Жуковского, вып. 335, 1949). — *Прим. ред.*

Вне зависимости от фильтров, которые мы можем вводить в систему, выходная переменная всегда будет содержать в качестве составляющей колебание с частотой ω_0. Полезно отметить, что амплитуду этой составляющей иногда даже нежелательно снижать при фильтрации до величины, меньшей определенного уровня. В самом деле, такое колебание создает «динамическую



Фиг. 47

смазку», приводящую к уменьшению влияния сухого трения, зазоров в механических передачах и других паразитных нелинейных явлений, обычно ухудшающих качество следящих систем.

Выше мы не сказали ничего вполне определенного о способе введения в реле колебания $E_0 \sin \omega_0 t$. Мы только вскользь упомянули, что его может создавать вспомогательный генератор. Системы, в которых это колебание вводится именно таким образом, имеют известные преимущества в смысле гибкости работы. Но недостаток их состоит в излишнем усложнении оборудования. Мы дадим теперь краткое описание различных следящих систем колебательного управления, устроенных таким образом, что это колебание создает сама система.

Рассмотрим систему, представленную схематически на фиг. 47. Пусть система рассчитана так, что при отсутствии входного сигнала она колеблется с частотой ω_0, определяемой фазовыми сдвигами линейных элементов в цепи обратной связи. Как мы уже отмечали, реле ведет себя по отношению к этим колебаниям как подлинно линейное звено, обладающее частотной характеристикой, убывающей с ростом амплитуды колебаний. Амплитуда колебаний сама устанавливается так, что коэффициент

усиления по контуру, определяемый усилением реле и линейных звеньев, оказывается равным единице.

Предположим теперь, что на систему действует входной сигнал. Если соответствующий сигнал рассогласования на входе реле достаточно мал, то усиление постоянного колебания, осуществляемое этим релейным устройством, почти не меняется и система продолжает колебаться с частотой и амплитудой, значения которых почти не отличаются от первоначальных. Как уже было отмечено, по отношению к этим сигналам релейное устройство ведет себя как линейное звено, у которого коэффициент усиления на 6 *дб* меньше коэффициента усиления для постоянного колебания. Ясно, что при таких условиях данная следящая система является следящей системой колебательного управления, которая была описана выше. Единственным новым обстоятельством здесь является то, что частота и амплитуда модулирующего колебания $E_0 \sin \omega_0 t$ определяются самой системой, а не независимо от нее, как до сих пор молчаливо предполагалось [1]).

Во всех рассуждениях об этой системе как о следящей системе надо только приписать релейному устройству собственную эффективную передаточную функцию, а затем следовать методам, указанным в предыдущих главах [2]). Нам нет надобности в явной форме учитывать постоянное колебание. Однако требование, в силу которого система должна действовать еще и как генератор, в известной мере ограничивает наши возможности в смысле улучшения качества процесса отслеживания. Остановимся на этих ограничениях несколько подробнее.

[1]) Это свойство является основным в автоколебательных системах, к которым относится рассматриваемая система. Существование автоколебаний обеспечивается наличием в системе релейного звена [1, 6, 13, 42, 43]. — *Прим. перев.*

[2]) Физически линеаризируя систему посредством модуляции указанным здесь методом, мы можем рассматривать ее как линейную только с точки зрения процесса преобразования входного сигнала в выходной сигнал системы. Но по своим структурным свойствам система благодаря наличию реле остается нелинейной и поэтому работает в режиме автоколебаний, используемых для приближенной физической линеаризации процесса прохождения сигналов через систему. — *Прим. перев.*

Пусть $F(s)$ означает передаточную функцию цепи обратной связи для сигналов, определяемых равенством (6.14). Тогда передаточная функция для установившегося (модулирующего) колебания равна $2F(s)$, и ввиду того, что система всегда работает в колебательном режиме, передаточная функция $1+[1/2F(s)]$ замкнутой системы имеет чисто мнимый нуль $s=i\omega_0$. Поэтому

$$2F(i\omega_0) = -1.$$

Следовательно, кривая на диаграмме Найквиста, описываемая точкой $1/F(s)$ при перемещении точки s вдоль мнимой оси, необходимо проходит через точку -2. С другой стороны, для обеспечения удовлетворительного качества слежения течение этой кривой должно отвечать условиям, выявленным в гл. IV, включая и то условие, что она не должна проходить вблизи точки -1. Очевидно, что для рассматриваемых систем выполнение этих условий затруднительно в большей мере, чем в других случаях, где указанное ограничение отсутствует. В этом смысле такие автоколебательные следящие системы менее гибки, чем следящие системы колебательного управления, в которых модулирующее колебание создается независимым генератором.

При использовании автоколебательных систем необходимо соблюдать элементарную предосторожность: для того чтобы кривая, которая принудительно проходит через точку -2, избежала сближения с точкой -1, она должна пересекать вещественную ось в точке -2 перпендикулярно к последней. Для этого необходимо, чтобы в окрестности частоты колебаний системы функция $1/F(i\omega)$ изменялась медленно по величине и быстро по фазе.

6.6. Обобщенные следящие системы колебательного управления. Реле представляет собой нелинейное устройство. Но путем сочетания управляющего сигнала с синусоидальным колебанием высокой частоты и большой амплитуды можно добиться того, чтобы выходная переменная системы изменялась прямо пропорционально вход-

ной переменной[1]). Таким образом, основная идея следящих систем колебательного управления заключается в физической линеаризации нелинейной системы[2]). Ж. М. Лёб[3])[4]) показал, что эта идея применима к любой нелинейной системе, и назвал этот метод обобщенным методом физической линеаризации следящих систем. Полученные в результате применения этого метода системы мы будем называть обобщенными следящими системами колебательного управления.

Рассмотрим некоторую функцию $y(x)$, где y — выход системы, а x — ее вход. Если вместо переменной x мы подставим сумму $x+\varepsilon$, где ε гораздо меньше x по абсолютной величине, то, коль скоро $y(x)$ — функция голоморфная, можно разложить $y(x+\varepsilon)$ в ряд Тэйлора

$$y(x+\varepsilon)=y(x)+\varepsilon\left(\frac{dy}{dx}\right)_x+\varepsilon^2\frac{1}{2}\left(\frac{d^2y}{dx^2}\right)_x+\ldots . \quad (6.15)$$

Определим теперь вход x как периодическую функцию времени t с периодом T, считая ε постоянной. Тогда очевидно, что $y(x)$ также является периодической функцией времени с тем же периодом T. Тот же вывод справедлив и для функций dy/dx и d^2y/dx^2. Периодические же функции мы можем разлагать в ряды Фурье. Тогда, пренебрегая членами, содержащими ε в степенях выше первой, получим

$$\begin{aligned} y(x+\varepsilon)\approx a_{00}+\sum_{n=1}^{\infty}(a_{0n}\cos n\omega t+b_{0n}\sin n\omega t)+ \\ +\varepsilon\left[a_{10}+\sum_{n=1}^{\infty}(a_{1n}\cos n\omega t+b_{1n}\sin n\omega t)\right], \quad (6.16)\end{aligned}$$

где $\omega=2\pi/T$ — частота входного колебания x.

Если ε не является в точности постоянной, а представляет собой медленно изменяющуюся функцию от t в том смысле, что ее основная частота гораздо меньше

---

[1]) Эта пропорциональность, т. е. линейность, является приблизительной и будет тем более точной, чем меньше влияние высших гармоник. — *Прим. перев.*

[2]) См. примечание[2]) на стр. 129. — *Прим. перев.*

[3]) Loeb J. M., Ann. de Télécommunications, 5, 65—71 (1950).

[4]) См. также статью Лёба в [44]. — *Прим. перев.*

частоты ω, то уравнение (6.16) попрежнему остается приблизительно верным. Теперь будем рассматривать функцию $y(x)$ как соотношение между выходом и входом в нелинейном устройстве, функцию $\varepsilon(t)$ — как управляющий сигнал, а $x(t)$ — как наложенное колебание высокой частоты и большой амплитуды, не обязательно синусоидальное. Составляющая выхода нелинейного устройства, характеризующая влияние сигнала, определяется вторым членом в правой части (6.16). Так как частота ω гораздо выше частоты сигнала $\varepsilon(t)$, то периодическую функцию, изображаемую рядом Фурье

$$a_{10} + \sum_{n=1}^{\infty} (a_{1n} \cos n\omega t + b_{1n} \sin n\omega t),$$

можно рассматривать как несущую, тогда как $\varepsilon(t)$ представляет собой модулирующую функцию. Проведенное рассуждение основано на существовании непосредственной зависимости $y(x)$ между входом и выходом нелинейного устройства. Но, как показал Лёб[1]), уравнение (6.16) справедливо также и в случае более общей зависимости между y и x, определяемой посредством функционалов; это означает, что переменная y зависит в момент времени t не только от мгновенного значения x при том же t, но также от всех предшествовавших значений x. Это обобщенное понятие соотношения между входом и выходом включает в себя соотношения, определяемые характеристикой типа гистерезиса, в том числе и характеристиками зазоров в механических передачах, и пригодно для описания почти всех нелинейных устройств, встречающихся в прикладных задачах. Поэтому в обобщенных следящих системах колебательного управления выходным сигналом нелинейного звена служит модулированная несущая. Больше того, по отношению к сигналу вход и выход связаны линейной зависимостью.

Пусть теперь наложенное колебание имеет вид симметричной волны — например синусоиды или пилообразного колебания, изображенного на фиг. 48. При этом условии,

---

[1]) В работе, указанной в примечании[3]) на стр. 131. — *Прим. перев.*

если $y(x)$ – функция четная, т. е. если

$$\left.\begin{aligned} & y(x) = y(-x) \\ \text{и} \quad & \left(\frac{dy}{dx}\right)_x = -\left(\frac{dy}{dx}\right)_{-x}, \end{aligned}\right\} \tag{6.17}$$

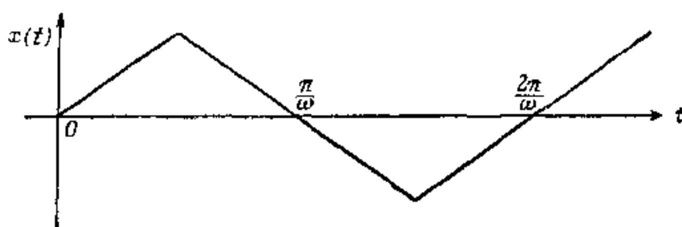
то между периодическими функциями $y(x)$ и dy/dx имеют место следующие фазовые соотношения:

$$\left.\begin{aligned} & y(x)_t = y(x)_{t+\frac{T}{2}}, \\ & \left(\frac{dy}{dx}\right)_t = -\left(\frac{dy}{dx}\right)_{t+\frac{T}{2}}. \end{aligned}\right\} \tag{6.18}$$

Из этих соотношений следует, что в данном случае

$$a_{01} = b_{01} = 0, \qquad a_{10} = 0 \qquad (y \text{ — четное}). \tag{6.19}$$

Поэтому, пренебрегая высшими гармониками, можно представить несущую как синусоидальное колебание частоты ω.



Фиг. 48

Но именно это и характеризует следящие системы на переменном токе, рассмотренные в предыдущих пунктах. Следовательно, описанный в них метод синтеза можно применить и к этому классу обобщенных следящих систем колебательного управления. Если же $y(x)$ — функция нечетная, т. е. если $y(x) = -y(-x)$, то для этого случая можно выписать равенства, аналогичные равенствам (6.17) и (6.18), и, далее,

$$a_{00} = 0, \qquad a_{11} = b_{11} = 0 \qquad (y \text{ — нечетное}). \tag{6.20}$$

Пренебрегая высшими гармониками, мы возвращаемся к системе, подобной следящей системе колебательного управления, рассмотренной в п. 6.4.

Из проведенного рассуждения следует вывод о том, что методом наложения установившегося колебания на входной сигнал можно линеаризировать характеристику нелинейного звена следящей системы и тем самым превратить систему в следящую систему колебательного управления улучшенного качества. Кроме того, такие следящие системы можно разрабатывать с помощью методов, уже изложенных в этой главе.

Глава VII

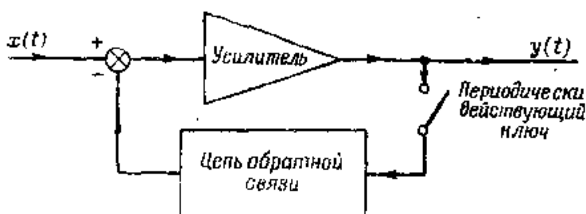
СЛЕДЯЩИЕ СИСТЕМЫ ПРЕРЫВИСТОГО ДЕЙСТВИЯ

Все рассмотренные до сих пор следящие системы предназначены для работы от сигналов, заданных в форме функции от непрерывной переменной t. Встречаются, однако, такие случаи, когда сигналы, от которых работает следящая система, задаются в форме функций от дискретной переменной. Эти случаи имеют место, например, в условиях, когда нам задан входной сигнал, полученный посредством определения значений функции $x(t)$ в равноотстоящие один от другого моменты времени $0, t_0, 2t_0, \ldots$. В этих условиях входной сигнал вовсе не определяется в интервалах времени между последовательными моментами съема сигнала.

Естественно, что в рассмотренных условиях нас интересуют значения выходного сигнала в моменты съема входного сигнала. Следовательно, следящая система должна действовать так, чтобы корректирующее воздействие на выходной сигнал зависело только от этих значений, а не от тех, которые выходной сигнал может принимать в интервалах вне точек съема. Следящую систему, предназначенную действовать указанным образом, можно назвать следящей системой прерывистого действия. В этой главе мы кратко изложим теорию линейных следящих систем указанного типа, которая, по существу, весьма сходна с теорией непрерывных следящих систем, разобранной в предыдущих главах.

7.1. Выход цепи прерывистого действия. В качестве простейшего образца следящей системы прерывистого действия мы будем рассматривать систему, изображенную на фиг. 49. Эта система содержит обычные цепи: прямую цепь и цепь обратной связи. Существенно новая особен-

ность системы состоит в наличии в цепи обратной связи ключа, действующего периодически так, что эта цепь бывает замкнута только на протяжении коротких промежутков времени, в равноотстоящие один от другого моменты времени 0, t_0, $2t_0$, ...[1]). Размещение внутри



Фиг. 49

системы звеньев, запасающих энергию (частотно-селективных звеньев), отражается на соответственной теории только в деталях. Мы воспользуемся поэтому возможностью несколько упростить изложение, допустив, что передаточная функция прямой цепи не зависит от частоты[2]). В этом случае существенно, чтобы ключ был расположен в указанном на фигуре месте.

Дальнейшее исследование основано на предположении о том, что промежутки времени, на протяжении которых ключ остается замкнутым, достаточно коротки; поэтому можно считать, что на цепь обратной связи воздействуют последовательные импульсы и что реакция $h_2(t)$ цепи обратной связи на импульс является непрерывной функцией. Последнее предположение означает, что при малых значениях t реакция $h_2(t)$ имеет порядок роста t^n, $n \gg 1$. Тогда $h_2(t)$ не имеет разрыва при $t=0$. Обозначив через

---

[1]) Теоретически эти короткие промежутки времени суть изолированные моменты, т. е. точки на оси t, почему и можно говорить об этих промежутках как о моментах 0, t_0, ... — *Прим. перев.*

[2]) «Как будет видно из дальнейшего, вопреки утверждению автора, допущение о постоянстве передаточной функции прямой цепи лежит в основе всего последующего анализа. При отказе от этого допущения уравнение (7.2) становится неприменимым и, по существу, мы приходим к значительно более сложной задаче, требующей иного метода решения» (примечание Я. З. Цыпкина на стр. 108 книги Маккола (см. стр. 116), материал которой используется Цяном в этой главе). — *Прим. перев.*

$F_2(s)$ передаточную функцию цепи обратной связи, мы установим, что, в силу общего соотношения (2.18),

$$h_2(t) = \frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} F_2(s) e^{ts} \, ds, \tag{7.1}$$

где γ — вещественная постоянная, превышающая вещественные части всех полюсов функции $F_2(s)$. Если при достаточно больших s функция $F_2(s)$ имеет порядок роста $1/s^m$, то на основании «словаря» оригиналов и изображений (табл. 1, стр. 23) функция $h_2(t)$ имеет при малых t порядок роста t^{m-1}. Из условия непрерывности функции $h_2(t)$ при $t = 0$ следует, что m должно равняться по крайней мере двум. Таким образом, при достаточно больших s, $F_2(s)$ стремится к нулю по крайней мере так же быстро, как функция $1/s^2$.

Теперь уже очевидно, что если входной сигнал $x(t)$ тождественно обращается в нуль при отрицательных значениях t, то значение выходного сигнала $y(t)$ в выбранный момент nt_0 съема сигнала определяется суммой влияний всех предшествующих импульсов и выражается формулой

$$y(nt_0) = F_1 \left[x(nt_0) - \theta t_0 \sum_{k=0}^{n} y(kt_0) h_2(nt_0 - kt_0) \right], \tag{7.2}$$

где F_1 — передаточная функция прямой цепи, а θ — доля периода включения, на протяжении которой ключ замыкает цепь. Произведение $\theta t_0 y(kt_0)$ равно, таким образом, импульсу, поступающему на вход цепи обратной связи в момент $t = kt_0$.

Если $x(t)$ и $h_2(t)$ известны в моменты съема сигнала, то значения $y(0)$, $y(t_0)$, $y(2t_0)$, ... можно последовательно вычислить элементарным путем с помощью формулы (7.2). Однако вместо того, чтобы идти этим путем, мы выберем более ясный путь, который позволит установить связь между теорией следящих систем прерывистого действия и теорией обычных следящих систем, разобранной в главе IV. Такой подход принадлежит Г. Р. Штибицу и К. Э. Шэннону[1]).

---

[1]) См. также книгу Я. З. Цыпкина [50]. — *Прим. ред.*

7.2. Теория Штибица и Шэннона. Введем функции

$$X^*(s) = \sum_{n=0}^{\infty} x(nt_0)\, e^{-nt_0 s}, \tag{7.3}$$

$$Y^*(s) = \sum_{n=0}^{\infty} y(nt_0)\, e^{-nt_0 s}, \tag{7.4}$$

$$F_2^*(s) = t_0 \sum_{n=0}^{\infty} h_2(nt_0)\, e^{-nt_0 s}. \tag{7.5}$$

Эти функции являются периодическими функциями комплексной переменной s с мнимым периодом $2\pi i/t_0$. Поэтому функции x, y и h_2, зависящие от аргумента nt_0, представляют собой коэффициенты Фурье. Техника перехода от функций $x(t)$, $y(t)$ и $h_2(t)$ к функциям $X^*(s)$, $Y^*(s)$ и $F_2^*(s)$ имеет очень много общего с техникой образования соответствующих изображений по Лапласу $X(s)$, $Y(s)$ и $F_2(s)$, определяемых соотношением (2.1). Здесь вместо непрерывно изменяющегося времени фигурируют дискретные значения времени nt_0 и, следовательно, знак интеграла заменяется знаком суммирования. Поэтому преобразования (7.3) – (7.5) представляют собой естественное видоизменение методики преобразования Лапласа при переходе к задачам, связанным со следящими системами прерывистого действия.

Сосредоточим пока свое внимание на том случае, когда все полюсы функции $F_2(s)$ расположены слева от мнимой оси. Тогда функция $h_2(t)$ при $t \to \infty$ убывает по экспоненциальному закону, и ряд в правой части равенства (7.5) сходится при всех значениях s, у которых вещественные части превышают некоторую отрицательную постоянную.

Что касается рядов в правых частях (7.3) и (7.4), то их сходимость зависит, конечно, от характера входного сигнала. Мы ограничимся рассмотрением только тех входных сигналов, при которых эти ряды сходятся таким же образом, как и ряд (7.5). Это сводится лишь к незначительным ограничениям, налагаемым на входные сигналы $x(t)$ и обычно принимаемым в теории неустановившегося режима.

Умножая все члены в обеих частях равенства (7.2) на e^{-nt_0s} и затем суммируя по всем значениям n, получаем

$$Y^*(s) = F_1\left[X^*(s) - \theta t_0 \sum_{n=0}^{\infty} e^{-nt_0s} \sum_{k=0}^{n} y(kt_0)\, h_2(nt_0 - kt_0)\right].$$

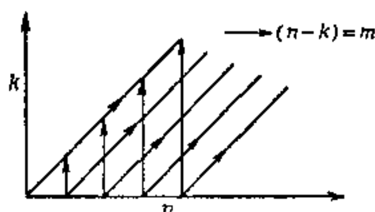
Но

$$\theta t_0 \sum_{n=0}^{\infty} e^{-nt_0s} \sum_{k=0}^{n} y(kt_0)\, h_2(nt_0 - kt_0) =$$

$$= \theta t_0 \sum_{n=0}^{\infty} \sum_{k=0}^{n} y(kt_0)\, e^{-kt_0s} e^{-(n-k)t_0s} h_2(nt_0 - kt_0) =$$

$$= \theta t_0 \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} y(kt_0)\, e^{-kt_0s} e^{-mt_0s} h_2(mt_0) = \theta F_2^*(s)\, Y^*(s).$$

Процесс перехода от суммирования по n и k к суммированию по k и $m = n - k$ графически пояснен на фиг. 50,



Фиг. 50

где заштрихованная область является областью, в которой выполняется суммирование. Поэтому

$$Y^*(s) = F_1\left[X^*(s) - \theta F_2^*(s)\, Y^*(s)\right]$$

или

$$\frac{Y^*(s)}{X^*(s)} = \frac{F_1}{1 + \theta F_1 F_2^*(s)}\,. \qquad (7.6)$$

Сотношение (7.6) является аналогом основного соотношения (4.3), относящегося к следящим системам непрерывного действия, рассмотренным выше. Различие же между этими соотношениями обусловлено свойствами соответствующих функций. Мы остановимся на этом позже

Примем теперь такое предположение о характере функции $y(nt_0)$, в силу которого ряд, определяющий функцию

$Y^*(s)$, сходится при всех значениях s, обладающих неотрицательными вещественными частями. В этом случае данный ряд сходится и при чисто мнимых значениях s. Пусть $s = i\omega$. Тогда

$$Y^*(i\omega)\,e^{int_0\omega} = \sum_{m=0}^{\infty} y(mt_0)\,e^{-it_0\omega(m-n)}.$$

Поэтому

$$\int_{\omega_0}^{\omega_0+\frac{2\pi}{t_0}} Y^*(i\omega)\,e^{int_0\omega}\,d\omega = y(nt_0)\,\frac{2\pi}{t_0}$$

или

$$y(nt_0) = \frac{t_0}{2\pi i}\int_{\omega_0}^{\omega_0+\frac{2\pi}{t_0}} Y^*(i\omega)\,e^{int_0\omega}i\,d\omega.$$

Положив $i\omega = s$ и $i\omega_0 = s_0$, получим окончательно

$$y(nt_0) = \frac{t_0}{2\pi i}\int_{s_0}^{s_0+\frac{2\pi i}{t_0}} Y^*(s)\,e^{nt_0 s}\,ds.$$

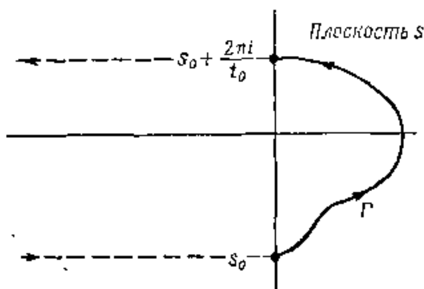
В силу теоремы Коши об интеграле от функции комплексной переменной имеем

$$y(nt_0) = \frac{t_0}{2\pi i}\int_{\Gamma} Y^*(s)\,e^{nt_0 s}\,ds = \frac{t_0}{2\pi i}\int_{\Gamma}\frac{F_1X^*(s)\,e^{nt_0 s}}{1+bF_1F_2^*(s)}\,ds, \qquad (7.7)$$

где Γ (фиг. 51) означает контур интегрирования, соединяющий две точки, удаленные одна от другой на расстояние $2\pi/t_0$ и лежащие на мнимой оси плоскости комплексной переменной s; этот контур проходит справа от всех особых точек подинтегральной функции. Приведенное здесь определение контура Γ является настолько общим, что уравнение (7.7) остается справедливым даже в том случае, когда функция $Y^*(s)$ имеет полюсы с положительной вещественной частью.

Благодаря периодичности функций $X^*(s)$ и $F_2^*(s)$ мы можем дополнить контур Γ пунктирными прямыми, парал-

лельными вещественной оси; значение интеграла при этом не изменится. Образованный таким образом контур будет содержать все полюсы подинтегральной функции. Но можно показать, что для широкого класса практически применимых входных функций функция $X^*(s)$ не будет иметь полюсов с положительными вещественными частями. В этом случае единственным возможным источником неустойчивости явится наличие в знаменателе подинтегральной функции в (7.7) нуля с положительной вещественной частью. Таким образом, совершенно аналогично условию устойчивости в следящих системах непрерывного действия необходимое и достаточное условие устойчивости следящих систем прерывистого действия заключается в том, чтобы уравнение



Фиг. 51

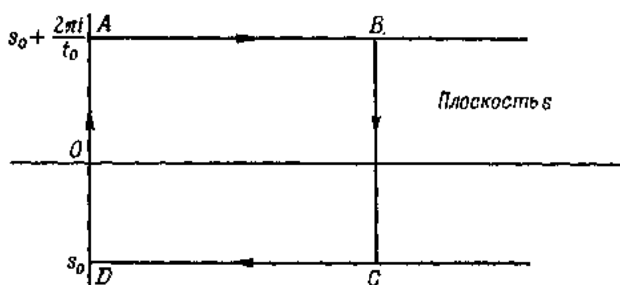
$$1+\theta F_1 F_2^*(s)=0 \qquad (7.8)$$

не имело корней в правой части плоскости комплексной переменной s. Покажем теперь, каким образом можно определить с помощью надлежащего видоизменения критерия Найквиста (п. 4.3), выполняется ли это условие.

7.3. Критерий Найквиста для следящих систем прерывистого действия. Ввиду периодичности функции $F_2^*(s)$ достаточно определить, имеет ли уравнение (7.8) корни в горизонтальной полуполосе шириной $2\pi/t_0$, простирающейся вправо от мнимой оси (фиг. 52), или же нет.

Предположим, что функция $F_2^*(s)$ не имеет особых точек на мнимой оси и вправо от этой оси, а также, что функция $1+\theta F_1 F_2^*(s)$ не имеет нулей на мнимой оси. Мы будем также предполагать, что упомянутая горизонтальная полуполоса расположена так, что функция $1+\theta F_1 F_2^*(s)$ не имеет нулей на ее горизонтальных сторонах. Пусть теперь точка s описывает замкнутую кривую $ABCDA$, показанную на

фиг. 52. Конец вектора $\theta F_1 F_2^*(s)$ описывает при этом некоторую замкнутую кривую типа кривой $A'B'C'D'A'$, изображенной на фиг. 53 (мы не пытаемся точно нанести кривую, описываемую концом этого вектора). Когда точка s пробегает отрезок AB, конец вектора перемещается вдоль кривой $A'B'$. Когда точка s пробегает отрезок BC, конец вектора перемещается вдоль дуги $B'C'$, причем вследствие



Фиг. 52

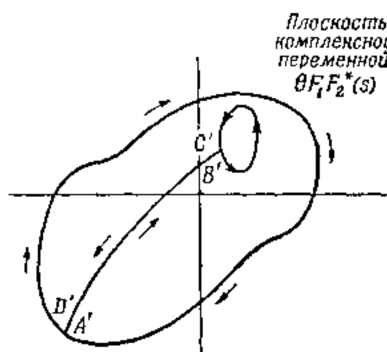
периодичности функции $F_2^*(s)$ $B'C'$ представляет собой замкнутую кривую. Когда точка s пробегает отрезок CD, конец упомянутого вектора перемещается вдоль дуги $C'D'$; вследствие периодичности функции $F_2^*(s)$ $C'D'$ совпадает с $A'B'$ с точностью до направления обхода. Наконец, когда точка s пробегает отрезок DA, конец вектора перемещается вдоль дуги $D'A'$, являющейся замкнутой кривой.

Согласно теореме Коши, уравнение (7.8) имеет корни в прямоугольнике $ABCD$ или не имеет их в этом прямоугольнике в зависимости от того, не равно или равно нулю число полных оборотов радиуса-вектора, проведенного из точки -1 к текущей точке $\theta F_1 F_2^*(s)$ при обходе этой текущей точкой контура $A'B'C'D'A'$.

Рассмотрим теперь, что происходит, когда сторона BC прямоугольника, показанного на фиг. 52, удаляется в бесконечность. С помощью соотношения (7.5) легко установить, что замкнутая кривая, образованная дугой $B'C'$ на фиг. 53, стягивается при этом в точку. Следовательно, характерная кривая сводится к дуге $D'A'$. Уравнение (7.8), очевидно, имеет корни в рассматриваемой полосе или не имеет их в ней в зависимости от того, отлично от нуля

или равно нулю число полных оборотов радиуса-вектора, проведенного из точки -1 к полученной таким образом предельной кривой. В этом и заключается критерий Найквиста для простых следящих систем прерывистого действия[1]).

Мы изложили здесь основы теории Штибица—Шэннона, посвященной исследованию следящих систем прерывистого действия; эта теория как по ее содержанию, так и по форме весьма сходна с теорией следящих систем непрерывного действия.



Фиг. 53

7.4. Установившаяся ошибка следящей системы прерывистого действия. Если на вход системы подана единичная функция, т. е.

$$X^*(s) = \sum_{n=0}^{\infty} e^{-nt_0 s} = \sum_{n=0}^{\infty} (e^{-t_0 s})^n = \frac{1}{1-e^{-t_0 s}},$$

то, согласно (7.7),

$$y(nt_0) = \frac{t_0 F_1}{2\pi i} \int\limits_{\Gamma} \frac{e^{nt_0 s}\,ds}{[1-e^{-t_0 s}][1+\theta F_1 F_2^*(s)]}.$$

При $n \to \infty$ единственным полюсом, имеющим существенное значение, является лишь полюс в начале и

$$\lim_{n\to\infty} y(nt_0) = \frac{F_1}{1+\theta F_1 F_2^*(0)}. \qquad (7.9)$$

Это соотношение определяет установившееся значение выхода при постоянном единичном входе. Поэтому условие

---

[1]) См. [5].—*Прим. перев.*

достаточной малости (по абсолютной величине) установившейся ошибки следящей системы прерывистого действия выражается соотношением

$$F_1 \approx 1 + \theta F_1 F_2^*(0)$$

или

$$F_1 \approx \frac{1}{1 - \theta F_2^*(0)}. \tag{7.10}$$

Последнее соотношение доставляет приближенное значение коэффициента усиления прямой цепи, которое будет иметь место при точном воспроизведении входа выходом. Соотношение (7.10) для следящих систем прерывистого действия аналогично соотношению (4.11) для следящих систем непрерывного действия.

7.5. Вычисление функции $F_2^*(s)$. Нам остается сделать еще один шаг, прежде чем можно будет сказать, что наша теория следящих систем прерывистого действия имеет практическое значение. Обе функции $F_2(s)$ и $F_2^*(s)$ являются характеристиками цепи обратной связи: $F_2(s)$ характеризует эту цепь, когда она применяется как часть следящей системы непрерывного действия, а $F_2^*(s)$ — когда цепь образует часть следящей системы прерывистого действия. С математической точки зрения первая функция гораздо проще второй; к тому же она является функцией, непосредственно используемой в известных методах расчета цепей. Поэтому важно установить возможно более прямую связь между этими функциями. С этой целью напомним, что $F_2^*(s)$ — функция, периодическая относительно s с чисто мнимым периодом $i2\pi/t_0$. Далее отметим, что при исследовании данной системы по способу Найквиста требуется рассматривать частотную характеристику $F^*(i\omega)$ только на интервале $-\frac{\pi}{t_0} < \omega < \frac{\pi}{t_0}$.

Предполагая, что вещественная часть s превосходит число γ, являющееся, в силу наших предположений,

отрицательным, подсчитаем $F_2^*(s)$ с помощью (7.1) и (7.5):

$$F_2^*(s)=\frac{t_0}{2\pi i}\sum_{n=0}^{\infty}e^{-nt_0s}\int_{\gamma-i\infty}^{\gamma+i\infty}F_2(q)\,e^{nt_0q}\,dq=$$

$$=\frac{t_0}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}F_2(q)\,dq\sum_{n=0}^{\infty}e^{-nt_0(s-q)}=$$

$$=\frac{t_0}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}\frac{F_2(q)\,dq}{1-e^{-t_0(s-q)}}. \qquad (7.11)$$

Вычислим правую часть равенства (7.11) методом вычетов. Полюсами подинтегральной функции являются полюсы функции $F_2(s)$, лежащие слева от пути интегрирования, и корни уравнения $1-e^{-t_0(s-q)}=0$, лежащие справа от него. Легко видеть, что интегрирование по прямой от $\gamma-i\infty$ до $\gamma+i\infty$ равносильно интегрированию по часовой стрелке по замкнутому контуру, образованному этой прямой и полуокружностью бесконечного радиуса, расположенной в правой полуплоскости комплексной переменной s. Следовательно, правая часть равенства (7.11) равна сумме вычетов подинтегрального выражения относительно различных корней уравнения $1-e^{-t_0(s-q)}=0$, умноженной на $-t_0$.

Но корнями последнего уравнения служат числа $q=s+(2\pi im/t_0)$, где m — целое число; вычет же подинтегральной функции относительно полюса, соответствующего этому корню, равен $-(1/t_0)F_2\left(s+\frac{2\pi im}{t_0}\right)$. Поэтому в результате получаем соотношение

$$F_2^*(s)=\sum_{m=-\infty}^{m=\infty}F_2\left(s+\frac{2\pi im}{t_0}\right). \qquad (7.12)$$

Эта формула позволяет более глубоко изучить свойства функции $F_2^*(s)$ и иногда оказывается полезной при приближенных подсчетах. Но можно легко получить точное выражение $F_2^*(s)$ в конечной форме.

Функцию $F_2(s)$ можно представить как сумму конечного числа элементарных дробей в виде

$$F_2(s) = \sum_{k=1}^{n} \frac{a_k}{s - s_k}, \qquad (7.13)$$

где все коэффициенты a_k и величины s_k — постоянные, а n — степень многочлена, служащего знаменателем функции $F_2(s)$. Подставив это выражение функции $F_2(s)$ в соотношение (7.12), получим

$$F_2^*(s) = \sum_{k=1}^{n} a_k \left\{ \frac{1}{s - s_k} + \sum_{m=1}^{\infty} \left[\frac{1}{\frac{2\pi im}{t_0} + (s - s_k)} - \frac{1}{\frac{2\pi im}{t_0} - (s - s_k)} \right] \right\} =$$

$$= \sum_{k=1}^{n} a_k \left[\frac{1}{s - s_k} + \sum_{m=1}^{\infty} \frac{2(s - s_k)}{\frac{4\pi^2 m^2}{t_0^2} + (s - s_k)^2} \right]. \qquad (7.14)$$

Выпишем теперь известное представление $\operatorname{cth} z$ с помощью ряда

$$\operatorname{cth} z = \frac{1}{z} + 2z \sum_{m=1}^{\infty} \frac{1}{m^2\pi^2 + z^2}.$$

Благодаря последнему соотношению можно фактически выполнить суммирование по индексу m; тогда

$$F_2^*(s) = \frac{t_0}{2} \sum_{k=1}^{n} a_k \operatorname{cth} \left[\frac{(s - s_k) t_0}{2} \right]. \qquad (7.15)$$

Эта формула позволяет точно вычислить $F_2^*(s)$ при любом значении s.

Когда t_0 очень мало и величиной $F_2(i\omega)$ можно пренебречь вне интервала $-\pi/t_0 < \omega < \pi/t_0$, качественное течение функции $F_2^*(i\omega)$ непосредственно очевидно из соотношения (7.12). В самом деле, в интервале $-\pi/t_0 < \omega < \pi/t_0$

функция $F_2^*(i\omega)$ приближенно равна функции $F_2(i\omega)$. Точно также мы убедимся, что, когда t_0 велико, для $F_2^*(i\omega)$ можно получить столь же простое приближение. Для этого запишем корень s_k в виде

$$s_k = -\lambda_k + i\omega_k, \tag{7.16}$$

где все λ_k и ω_k — вещественные величины. Согласно принятым допущениям, все величины λ_k положительны. Тогда соотношение (7.15) можно преобразовать следующим образом:

$$F_2^*(i\omega) = \frac{t_0}{2} \sum_{k=1}^{n} a_k \operatorname{cth} \left\{ \frac{t_0}{2} [\lambda_k + i(\omega - \omega_k)] \right\} =$$

$$= \frac{t_0}{2} \sum_{k=1}^{n} a_k \frac{1 + e^{-t_0[\lambda_k + i(\omega - \omega_k)]}}{1 - e^{-t_0[\lambda_k + i(\omega - \omega_k)]}}.$$

Поэтому, когда t_0 велико,

$$F_2^*(i\omega) \approx \frac{t_0}{2} \sum_{k=1}^{n} a_k \{1 + 2e^{-t_0[\lambda_k + i(\omega - \omega_k)]}\}. \tag{7.17}$$

Когда переменная s принимает большие значения, разложение (7.13) можно представить в виде

$$F_2(s) = \frac{1}{s} \sum_{k=1}^{n} a_k + \frac{1}{s^2} \sum_{k=1}^{n} a_k s_k + \ldots .$$

Но мы предположили, что $F_2(s)$ приближается к нулю по меньшей мере как $1/s^2$, когда s неограниченно возрастает (что обусловливает непрерывность выходной реакции цепи обратной связи на импульсное воздействие); следовательно,

$$\sum_{k=1}^{n} a_k = 0. \tag{7.18}$$

Поэтому равенство (7.17) преобразуется к виду

$$F_2^*(i\omega) \approx t_0 e^{-it_0\omega} \sum_{k=1}^{n} a_k e^{t_0 s_k}. \tag{7.19}$$

В реальных системах корни s_k или представляют собой вещественные числа, или же встречаются в виде сопряженных пар комплексных чисел. Поэтому конечная сумма в правой части (7.19) является всегда величиной вещественной и кривая, описываемая точкой $F_2^*(i\omega)$ при изменении ω от $-\pi/t_0$ до π/t_0, представляет собой окружность радиуса

$$t_0 \sum_{k=1}^{n} a_k e^{t_0 s_k} \Big|. \tag{7.20}$$

7.6. Сравнение следящих систем непрерывного и прерывистого действия. Мы убедились в том, что при малых значениях t_0 функция $F_2^*(i\omega)$ приблизительно совпадает с $F_2(i\omega)$. В следящих системах непрерывного действия критерий устойчивости заключается в том, чтобы график функции $F_1F_2(i\omega)$ не подходил к окрестности точки -1. В следящих системах прерывистого действия к этой окрестности не должен подходить график функции $\theta F_1F_2^*(i\omega)$ или, иначе, график функции $F_1F_2^*(i\omega)$ не должен подходить к окрестности точки $-1/\theta$. Поэтому если нас интересует только вопрос об устойчивости, то следящие системы прерывистого действия могут работать при гораздо больших значениях коэффициента усиления, чем системы непрерывного действия.

Когда значения t_0 достаточно велики, то, в силу соотношений (7.19) и (7.20), критерий Найквиста сводится к простому неравенству

$$\theta F_1 \Big| t_0 \sum_{k=1}^{n} a_k e^{t_0 s_k} \Big| < 1.$$

Так как вещественные части корней s_k отрицательны, то радиус кривой, описываемой точкой $F_2^*(i\omega)$, очень мал. Учитывая это обстоятельство, а также то, что величина θ, характеризующая продолжительность замыкания ключа, мала, мы приходим к следующему выводу: прямая цепь системы прерывистого действия может работать при очень высоких значениях коэффициента усиления без наступления неустойчивости. Таким образом, при любой продолжительности периода t_0 включения требования к следящей системе прерывистого действия, обусловливающие ее устойчивость,

налагают на систему гораздо меньшие ограничения по сравнению с соответствующими ограничениями для следящих систем непрерывного действия. Может быть, этого и следовало ожидать, так как промежутки времени, на протяжении которых цепь обратной связи оказывается замкнутой, весьма коротки, а на течение выходной переменной системы вне этих промежутков не налагается никаких воздействий.

7.7. Случай, когда функция $F_2(s)$ имеет полюс в начале. В практических условиях вполне возможно, что функция $F_2(s)$ имеет полюс при $s=0$. До сих пор мы исключали этот случай из рассмотрения, с тем чтобы избежать некоторых, впрочем, незначительных усложнений. Теперь мы его кратко рассмотрим.

Прежде всего заметим, что если величина $s=0$ является полюсом функции $F_2(s)$, то постоянная γ должна быть положительной, а наше представление этой функции в виде ряда останется справедливым только при значениях s, обладающих положительной вещественной частью. Далее отметим, что диаграмма Найквиста для этой системы также претерпевает изменения, а именно вместо замкнутой кривой мы получим разомкнутую кривую, концы которой надо считать соединенными полуокружностью бесконечного радиуса.

Однако полученное выше выражение для функции $F_2^*(s)$ в конечной форме остается в силе. Положив $s_1=0$, мы получим на основании равенства (7.15)

$$F_2^*(i\omega)=\frac{t_0}{2}\left[-ia_1\operatorname{ctg}\frac{\omega t_0}{2}+\sum_{k=2}^{n}a_k\operatorname{cth}\frac{t_0[\lambda_k+i(\omega-\omega_k)]}{2}\right].$$

Следовательно, когда t_0 велико, мы приходим к выражению, аналогичному (7.17), т. е. к выражению

$$F_2^*(i\omega)=\frac{t_0}{2}\left[-ia_1\operatorname{ctg}\frac{\omega t_0}{2}+\sum_{k=2}^{n}a_k\{1+2e^{-t_0[\lambda_k+i(\omega-\omega_k)]}\}\right].$$

Но, в силу равенства (7.18),

$$a_2+a_3+\ldots+a_n=-a_1$$

и поэтому

$$F_2^*(i\omega) = -\frac{a_1 t_0}{2}\left(1 + i \operatorname{ctg}\frac{\omega t_0}{2}\right) + t_0 e^{-it\omega} \sum_{k=2}^{n} a_k e^{t_0 s_k}. \qquad (7.21)$$

Постоянная a_1, конечно, вещественна и положительна. При изменении ω от $-\pi/t_0$ до $+\pi/t_0$ первое слагаемое в правой части (7.21) определяет прямую, параллельную мнимой оси. Второе слагаемое в выражении для $F_2^*(i\omega)$ определяет синусоидальную функцию. Отсюда следует, что кривая Найквиста представляет собой результат синусоидальной деформации прямой линии.

Глава VIII

ЛИНЕЙНЫЕ СИСТЕМЫ С ЗАПАЗДЫВАНИЕМ[1])

В настоящей главе мы введем в рассматриваемые нами линейные системы с постоянными коэффициентами еще один новый элемент: *запаздывание по времени*. Говоря о запаздывании τ по времени, мы имеем в виду, что соотношение между различными переменными системы нельзя выразить в виде такого соотношения между этими переменными, в котором значения всех переменных относятся к одному и тому же моменту времени t; наоборот, значения некоторых из переменных соответствуют моменту t, а значения некоторых других переменных — более раннему моменту $t-\tau$. Следовательно, переменные, заданные для момента $t-\tau$, запаздывают на промежуток времени τ по отношению к переменным, заданным для момента t. Таким образом, это запаздывание по времени коренным образом отличается от постоянной времени линейной системы первого порядка, рассмотренной в п. 3.1. Системы с запаздыванием по времени (с запаздывающим аргументом) описываются дифференциально-разностными уравнениями с постоянными коэффициентами и являются более сложными, чем линейные системы, рассматривавшиеся до сих пор и описывавшиеся дифференциальными уравнениями. Системы с запаздывающим аргументом изучали многие исследователи, в частности А. Каллендер, Д. Хартри и А. Портер[2]) и Н. Минорский[3]). Нас, однако, интересует несколько более ограниченная задача. Мы хотим знать,

---

[1]) См. также [13, 29]. — *Прим. перев.*

[2]) Callander A., Hartree D., Porter A., Trans. Roy. Soc., London (A), **235**, 415—444 (1935). (Изложение этой работы имеется в книге [13]. — *Прим. перев.*)

[3]) Minorsky N., Journ. Appl. Mech. (ASME), **9**, 67—71 (1942).

каким способом можно исследовать поведение следящей системы (с обратной связью), содержащей запаздывание τ по времени. В частности, мы хотим видоизменить способ Найквиста (п. 3.4) так, чтобы он стал применим к системам с запаздывающим аргументом.

Мы изложим теорию систем с запаздывающим аргументом, рассматривая конкретный пример такой системы, а именно пример, относящийся к стабилизации режима горения в ракетном двигателе с помощью системы с обратной связью. Вопрос о неустойчивости режима горения в ракетном двигателе изучали многие авторы. В нашем исследовании влияния запаздывания в ракетном двигателе мы будем следовать работе Л. Крокко[1]). Для простоты рассуждения[2]) рассмотрим только случай так называемых низкочастотных колебаний в ракетном двигателе, работающем на однокомпонентном жидком топливе.

8.1. Запаздывание во времени при горении. Введем обозначения: $\dot{m}_b(t)$ — масса горячего газа, образуемая при горении в единицу времени и соответствующая моменту времени t, $\dot{m}_i(t)$ — расход топлива на сгорание в момент t, $\tau(t)$ — запаздывание по времени (временно́е запаздывание) той частицы жидкого топлива, которая сгорает в момент t. Тогда масса топлива, сгоревшая за промежуток времени между моментами t и $t+dt$, должна равняться массе топлива, впрыснутой в камеру горения за промежуток времени между моментами $t-\tau$ и $t-\tau+d(t-\tau)$. Следовательно,

$$\dot{m}_b(t)\,dt = \dot{m}_i(t-\tau)\,d(t-\tau). \qquad (8.1)$$

Горячие газы, образовавшиеся в результате сгорания топлива, или используются для наполнения камеры горения для повышения их давления $p(t)$, или же выпускаются наружу через сопло ракеты в форме реактивной струи. Если частота возможных колебаний в камере го-

---

[1]) Crocco L., Journ. Amer. Rocket Soc., **21**, 163—178 (1951).

[2]) Последующее рассуждение основано на работе, опубликованной в Journ. Amer. Rocket Soc., **22**, 256—262 (1952).

рения низка, то давление в камере можно считать однородным и в первом приближении[1]) газовый поток через сопло можно рассматривать как квазистационарный. В этих предположениях массовый расход газа через сопло пропорционален плотности горячего газа в ракетном двигателе. Но в двигателе, работающем на однокомпонентном топливе, температура горячего газа почти не зависит от давления, при котором происходит сгорание топлива, и поэтому плотность этого газа пропорциональна только его давлению. Пусть $\overline{\dot{m}}$ — установившийся расход массы в системе, $\overline{M}_g$ — среднее значение массы газа, находящейся в двигателе, $\overline{p}$ — установившееся значение давления в камере горения. Пренебрегая объемом, занятым не сгоревшей еще частью жидкости, имеем

$$\dot{m}_b dt = \overline{\dot{m}} \left(\frac{p}{\overline{p}} \right) dt + d \left(\overline{M}_g \frac{p}{\overline{p}} \right). \tag{8.2}$$

Введем безразмерные переменные φ и μ для давления газа в камере горения и расхода топлива на сжигание, определяемые с помощью соотношений

$$\varphi = \frac{p - \overline{p}}{\overline{p}}, \qquad \mu = \frac{\dot{m}_b - \overline{\dot{m}}}{\overline{\dot{m}}}. \tag{8.3}$$

Следовательно, φ и μ представляют собой относительные отклонения соответственно давления газа и расхода топлива от их установившихся значений. Исключая $\dot{m}_b$ из соотношений (8.1) и (8.2) и учитывая (8.3), приходим к уравнению

$$\frac{\overline{M}_g}{\overline{\dot{m}}} \cdot \frac{d\varphi}{dt} + \varphi + 1 = \left(1 - \frac{d\tau}{dt} \right) [\mu (t - \tau) + 1]. \tag{8.4}$$

Для подсчета величины $d\tau/dt$ воспользуемся теорией Крокко о зависимости давления от запаздывания в сгорании. Пусть $f(p)$ — функция, определяющая скорость подготовки жидкого топлива к окончательному быстрому превращению в горячий газ; тогда запаздывание τ под-

[1]) Tsien H. S., Journ. Amer. Rocket Soc., 22, 139—143 (1952).

считывается с помощью соотношения

$$\int\limits_{t-\tau}^{t} f(p)\,dt = \text{const}. \qquad (8.5)$$

Постоянную в правой части (8.5) можно рассматривать как количество тепла, которое необходимо сообщить единице массы холодного впрыскиваемого топлива до ее воспламенения. Тогда функцию $f(p)$ можно физически истолковать как скорость теплопередачи от горячих газообразных продуктов горения к впрыскиваемому в камеру горения жидкому топливу. Дифференцируя обе части (8.5) по t, получаем

$$[f(p)]_t - [f(p)]_{t-\tau}\left(1 - \frac{d\tau}{dt}\right) = 0.$$

Теперь мы в состоянии дать в явной форме определение малого отклонения однородной системы от ее установившегося состояния. Допустим, что отклонение давления p от его установившегося значения невелико. Тогда функцию $f(p)$ можно разложить в ряд Тэйлора около $p=\bar{p}$ один раз для момента t, а другой раз — для момента $t-\tau$. Ограничиваясь учетом только членов порядков не выше первого, получим

$$[f(p)]_t = f(\bar{p}) + \bar{p}\left(\frac{df}{dp}\right)_{p=\bar{p}} \varphi(t),$$

$$[f(p)]_{t-\tau} = f(\bar{p}) + \bar{p}\left(\frac{df}{dp}\right)_{p=\bar{p}} \varphi(t-\tau).$$

Здесь τ — запаздывание при среднем значении $\bar{p}$ давления, т. е. теперь τ — постоянная. Следовательно,

$$1 - \frac{d\tau}{dt} = 1 - \left(\frac{d\ln f}{d\ln p}\right)_{p=\bar{p}} [\varphi(t) - \varphi(t-\tau)]. \qquad (8.6)$$

С помощью уравнений (8.4) и (8.6) придем к выражению

$$\frac{d\varphi}{dz} + \varphi = \mu(z-\delta) + n[\varphi(z) - \varphi(z-\delta)], \qquad (8.7)$$

где

$$n = \left(\frac{d\ln f}{d\ln p}\right)_{p=\bar{p}} \qquad (8.8)$$

и

$$\theta_g = \frac{\overline{M}_g}{\dot{m}}, \qquad z = \frac{t}{\theta_g}, \qquad \delta = \frac{\tau}{\theta_g}, \tag{8.9}$$

а постоянная θ_g равна отношению среднего значения газообразной массы, находящейся в двигателе, к среднему значению расхода массы через двигатель. Таким образом, θ_g равно среднему значению промежутка времени между моментом образования горячего газа в результате сгорания топлива и моментом его истечения из реактивного сопла. По этой причине θ_g называют постоянной времени газового потока. При дальнейших подсчетах мы будем выражать время в единицах этой основной постоянной времени. Переменная z представляет собой безразмерное время, а δ — безразмерную постоянную запаздывания процесса горения.

Функция $f(p)$ пропорциональна p^n, где n — постоянная, не зависящая от $\overline{p}$. В такой форме задает функцию $f(p)$ Крокко. Мы же рассмотрим случай, несколько более общий в том отношении, что зависимость $f(p)$ может быть произвольной, а величина n подсчитывается с помощью (8.8) и является функцией от $\overline{p}$. Если рассматривать $f(p)$ как скорость теплопередачи от горячих газообразных продуктов горения к впрыскиваемому в камеру горения жидкому топливу, то из физических законов теплопередачи следует, что значение n заключено между $\frac{1}{2}$ и 1.

8.2. Диаграмма Сатче. Неустойчивость процесса горения при равномерной подаче топлива в камеру горения Крокко назвал неустойчивостью, органически присущей этому процессу. Если количество топлива, подаваемого в единицу времени в камеру горения, определяется постоянной, не зависящей от давления p в этой камере, то $\mu \equiv 0$. Следовательно, устойчивость системы определяется простым уравнением, получаемым из уравнения (8.7):

$$\frac{d\varphi}{dz} + (1-n)\varphi(z) + n\varphi(z-\delta) = 0. \tag{8.10}$$

К уравнению (8.10) можно применить метод преобразования Лапласа так, как он применяется к уравнениям,

не содержащим запаздывающего аргумента и рассмотренным в предыдущих главах. Именно в этом заключается прием, который использовал Г. И. Ансов[1]). Однако данное основное уравнение, с помощью которого исследуется задача об устойчивости горения, не содержит вынуждающего члена. Поэтому применение классического метода решения линейных дифференциально-разностных уравнений доставляет более прямой подход к задаче. Таким образом, мы будем искать $\varphi(z)$ в виде экспоненты:

$$\varphi(z) \sim e^{sz}.$$

Подставив это значение $\varphi(z)$ в уравнение (8.10), получим характеристическое уравнение

$$s + (1-n) + ne^{-\delta s} = 0 \tag{8.11}$$

относительно показателя s степени.

Уравнение (8.11) можно также получить и с помощью преобразования Лапласа, применив последнее к исходному уравнению (8.10). Умножая обе части этого уравнения на e^{-sz} и затем интегрируя по z в пределах от $z=0$ до $z=\infty$ и обозначая через $\Phi(s)$ изображение по Лапласу функции $\varphi(z)$, приходим к равенству

$$s\Phi(s) - \varphi(0) + (1-n)\Phi(s) + n\int_0^\infty \varphi(z-\delta)e^{-sz}\,dz = 0.$$

Но

$$\int_0^\infty \varphi(z-\delta)e^{-sz}\,dz = e^{-s\delta}\int_0^\infty \varphi(z-\delta)e^{-s(z-\delta)}\,dz = e^{-s\delta}\left[\Phi(s) + \int_{-\delta}^{0} \varphi(z')e^{-sz'}\,dz'\right].$$

Поэтому если выбрать начальные значения так, что $\varphi = 0$ при $z \leqslant 0$ (так называемые нулевые начальные условия), то

$$[s + (1-n) + ne^{-s\delta}]\,\Phi(s) = 0,$$

---

[1]) Ansoff H. I., Journ. Appl. Mech. (ASME), 16, 158—164 (1949).

откуда мы вновь получаем уравнение (8.11). Здесь комплексная переменная s имеет тот же смысл, который она имела в предыдущих главах. Единственное различие между комплексными переменными s настоящей главы и предшествующих глав состоит в том, что здесь комплексная переменная s приведена к безразмерной форме с помощью постоянной θ_g времени газового потока. Интересно также отметить, что если бы уравнение (8.10) было неоднородным, т. е. содержало бы в правой части вынуждающий член, то уравнение, которое получилось бы в результате применения к упомянутому неоднородному уравнению преобразования Лапласа, оказалось бы также неоднородным. Очевидно, что функция

$$F(s) = \frac{1}{s + (1-n) + ne^{-\delta s}}$$

служит передаточной функцией рассматриваемой системы, причем выходной функцией служит $\varphi(z)$. Таким образом, функция $F(s)$ является еще одним примером трансцендентной передаточной функции.

Крокко нашел значение комплексного корня s с помощью решения системы двух уравнений, соответствующих вещественной и мнимой частям уравнения (8.11). Однако если нас интересует задача об исследовании устойчивости системы, то можно с успехом воспользоваться теоремой Коши, приведенной в п. 4.3. Введем обозначение

$$G(s) = e^{-\delta s} - \left(-\frac{1-n}{n} - \frac{s}{n}\right). \tag{8.12}$$

Тогда ответ на вопрос об устойчивости системы будет зависеть от того, имеет ли функция $G(s)$ нули в правой полуплоскости комплексной переменной s. А на вопрос о наличии или отсутствии таких нулей $G(s)$ можно ответить, если проследить за изменением аргумента функции $G(s)$, когда s обегает полуокружность, охватывающую правую полуплоскость s (фиг. 31). Если при этом вектор $G(s)$ совершает целое число оборотов по часовой стрелке, то на основании теоремы Коши это число оборотов равно разности между числом нулей и числом полюсов функции $G(s)$ в правой полуплоскости s. Так как, очевидно, $G(s)$ не имеет полюсов в этой правой полуплоскости, то

число оборотов вектора $G(s)$ равно числу нулей $G(s)$. Следовательно, условие устойчивости системы состоит в том, чтобы вектор $G(s)$ не совершал целого числа оборотов при обходе указанной полуокружности. Поэтому вопрос об устойчивости системы можно выяснить с помощью кривой Найквиста.

Однако непосредственное применение этого метода к исследованию интересующих нас свойств функции $G(s)$, определенной равенством (8.12), связано с известными неудобствами, обусловленными присутствием члена $e^{-\delta s}$, характеризующего запаздывание. М. Сатче[1]) предложил очень изящный и остроумный метод исследования таких систем с запаздывающим аргументом: вместо того чтобы рассматривать функцию $G(s)$ непосредственно, разобьем ее на две части:

$$G(s)=g_1(s)+g_2(s), \tag{8.13}$$

где

$$\begin{gathered} g_1(s)=e^{-\delta s}, \\ g_2(s)=-\frac{1-n}{n}-\frac{s}{n}. \end{gathered} \tag{8.14}$$

Таким образом, при этом разбиении конец вектора $G(s)$ лежит на кривой, описываемой концом вектора $g_1(s)$, а начало $G(s)$ лежит на кривой, описываемой концом вектора $g_2(s)$. Кривая $g_1(s)$ при перемещении s вдоль мнимой оси представляет собой окружность единичного радиуса. При значениях s, соответствующих точкам на полуокружности большого радиуса, точки $g_1(s)$ находятся внутри этого круга. Кривая $g_2(s)$ представляет собой прямую линию (фиг. 54), параллельную мнимой оси, когда s перемещается вдоль мнимой оси. Когда же s описывает полуокружность большого радиуса, то $g_2(s)$ представляет собой полуокружность большого радиуса, замыкающую упомянутую кривую слева. Нетрудно сообразить, что для того чтобы вектор $G(s)$ не совершал целого числа оборотов ни при каком значении временного запаздывания δ, кривая $g_2(s)$ должна лежать целиком вне кривой $g_1(s)$.

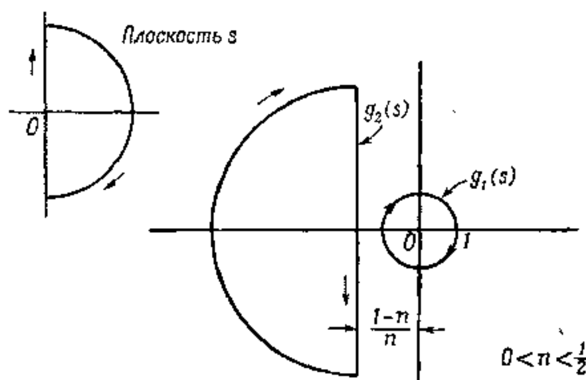
---

[1]) Satche M., Journ. Appl. Mech. (ASME), **16**, 419, 420 (1949).

Это означает, что условие абсолютной устойчивости[1]) системы выражается неравенством

$$\frac{1-n}{n} > 1 \quad \text{или} \quad \frac{1}{2} > n > 0. \tag{8.15}$$

Теперь легко видеть, что разделение $G(s)$ на две части $g_1(s)$ и $g_2(s)$ приводит к большому упрощению исследуемых кривых. Графическое построение, связанное с использованием функций $g_1(s)$ и $g_2(s)$, мы назовем *диаграммой Сатче*.



Фиг. 54

При $n > \frac{1}{2}$ кривые $g_1(s)$ и $g_2(s)$ пересекаются. Однако система все еще может сохранять устойчивость; для этого требуется, чтобы внутри единичного круга на фиг. 55 кривая $g_1(s)$ протекала правее кривой $g_2(s)$. Последнее условие выполняется вместе с неравенством

$$\cos\left(\delta\sqrt{2n-1}\right) > -\frac{1-n}{n},$$

т. е. если

$$\delta < \delta^*,$$

[1]) То есть условие устойчивости, имеющей место при любых δ.— *Прим. перев.*

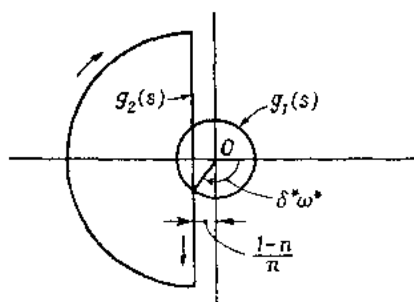
где

$$\delta^* = \frac{\arccos\left(-\frac{1-n}{n}\right)}{\sqrt{2n-1}} = \frac{1}{\sqrt{2n-1}}\left[\pi - \arccos\left(\frac{1-n}{n}\right)\right]. \quad (8.16)$$

Введем обозначение

$$\omega^* = \sqrt{2n-1}. \quad (8.17)$$

При $\delta = \delta^*$ имеет место равенство $G(i\omega^*) = 0$. Поэтому при $\delta = \delta^*$ уравнение (8.10) имеет колебательное решение по φ частоты ω^*. Таким образом, δ^* и ω^* представляют собой безразмерное критическое временнóе запаздывание и безразмерную критическую частоту.

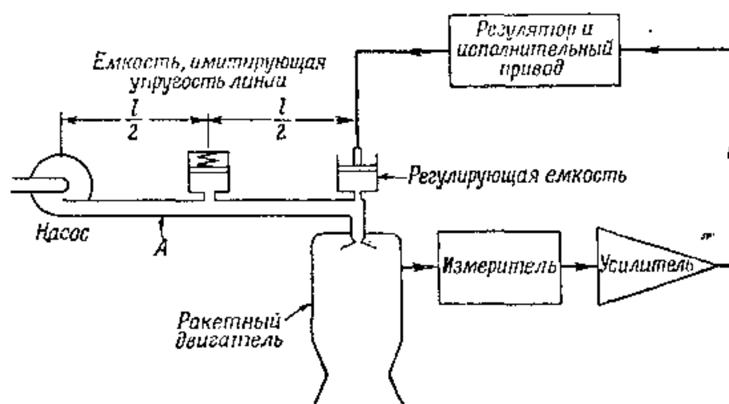


Фиг. 55

8.3. Динамика системы регулирования ракетного двигателя с обратной связью. Рассмотрим систему, схематически изображенную на фиг. 56 и состоящую из ракетного двигателя, питающего его насосного агрегата и цепи обратной связи. Имея целью приближенно учесть влияние упругости трубопровода, питающего двигатель, мы поместили в середине этого трубопровода емкость, имитирующую упругость линии, выполнив эту емкость в виде поршня, нагруженного пружиной. Вблизи форсунки расположена другая емкость, управляемая с помощью автоматического регулятора. Входом регулятора служит сигнал, снимаемый с датчика, измеряющего давление в камере горения; этот сигнал подается на вход регулятора после усиления. Коль скоро конструкции двигателя и насосного агрегата фиксированы, возникает вопрос о том, возможно ли спроектировать надлежащий усилитель таким образом, чтобы вся система в целом была устойчивой. При этом, так как мы не располагаем достоверными сведениями о величине временнóго запаздывания в процессе горения, в практических условиях следует исходить из необходимости обеспечения абсолютной устой-

чивости, т. е. устойчивости при любом значении запаздывания δ.

Обозначим через $\dot{m}_0$ мгновенный массовый расход топлива, подаваемого насосом, через p_0 — мгновенное значение давления топлива на выходе насоса, через $\bar{\dot{m}}$ — среднее значение расхода топлива и через $\bar{p}_0$ — значение давления.



Фиг. 56

Действие насоса описывается следующим уравнением:

$$\frac{p_0-\bar{p}_0}{\bar{p}_0}=-\alpha\frac{\dot{m}-\bar{\dot{m}}}{\bar{\dot{m}}}. \tag{8.18}$$

В условиях, когда скорость изменения массового расхода топлива мала по сравнению со скоростью распространения ударной волны в этом жидком топливе, но велика по сравнению со скоростью изменения числа оборотов ротора насоса, обычно небольшой коэффициент α представляет собой просто угловой коэффициент касательной к характеристической кривой насоса давление — объемный расход при *постоянном* числе оборотов ротора насоса, вычисленный в рабочей точке.

В обычных центробежных насосах α приблизительно равно единице. В насосах с постоянной производитель-

ностью α весьма велико. В насосах, работающих при постоянном давлении, или же в простых нагнетательных устройствах $\alpha = 0$.

Пусть $\dot{m}_1$ — мгновенный массовый расход топлива в части трубопровода, расположенной после емкости, имитирующей его упругие свойства, χ - упругая постоянная пружины, нагружающей поршень, а p_1 — мгновенное значение давления в области этой емкости. Тогда

$$\dot{m}_0 - \dot{m}_1 = \rho\chi \frac{dp_1}{dt}, \tag{8.19}$$

где ρ — плотность используемого жидкого топлива, $\rho = \text{const}$.

В последующих расчетах мы будем пренебрегать гидравлическими потерями в трубопроводе, обусловленными силами трения. В этом предположении падение $p_0 - p_1$ давления создается только за счет ускорения потока топлива, т. е.

$$p_0 - p_1 = \frac{l}{2A} \frac{d\dot{m}_0}{dt}, \tag{8.20}$$

где постоянная A равна площади поперечного сечения питающего трубопровода, а l — его полная длина. Подобным же образом, обозначив через p_2 мгновенное значение давления в области регулирующей емкости, напишем

$$p_1 - p_2 = \frac{l}{2A} \cdot \frac{d\dot{m}_1}{dt}. \tag{8.21}$$

Обозначив еще через C приращение емкости трубопровода (в единицах массы) при перемещении заслонки регулирующей емкости, будем иметь

$$\dot{m}_1 - \dot{m}_i = \frac{dC}{dt}. \tag{8.22}$$

Так как регулирующая емкость расположена очень близко к форсунке, инерцией части топлива, заполняющего объем между регулирующей емкостью и форсункой, можно пренебречь, и

$$p_2 - p = \frac{1}{2} \frac{\dot{m}_i^2}{\rho A_i^2}, \tag{8.23}$$

где A_i — эффективная площадь входного сечения форсунки. Величину A_i можно из расчета исключить, если заметить, что в установившемся состоянии разность $\Delta\overline{p}$ давлений p_0 и $\overline{p}$ определяется соотношением

$$\overline{p}_0 - \overline{p} = \Delta\overline{p} = \frac{1}{2}\frac{\dot{m}^2}{\rho A_i^2}. \tag{8.24}$$

Уравнения (8.18) — (8.24) описывают динамику системы подачи топлива. Непосредственное исключение из этих уравнений вспомогательных переменных позволяет получить соотношение, содержащее только переменные $\dot{m}_i$, p и C. При этом, чтобы выразить такое соотношение через безразмерные переменные, введем величины

$$P = \frac{\overline{p}}{2\Delta\overline{p}}, \qquad E = \frac{2\Delta\overline{p}}{\dot{\overline{m}}\theta_g}\rho\chi, \qquad J = \frac{l\dot{\overline{m}}}{2\Delta\overline{p}A\theta_g} \tag{8.25}$$

и

$$\varkappa = \frac{C}{\dot{\overline{m}}\theta_g}, \tag{8.26}$$

где θ_g — постоянная времени газового потока, определенная соотношением (8.9). В результате придем к искомому соотношению между безразмерными переменными φ, μ и $\varkappa$:

$$\begin{aligned}
P\left[1 + \alpha E\left(P + \frac{1}{2}\right)\frac{d}{dz} + \frac{1}{2}JE\frac{d^2}{dz^2}\right]\varphi +{}\\
+ \left\{\left[1 + \alpha\left(P + \frac{1}{2}\right)\right] + \left[\alpha E\left(P + \frac{1}{2}\right) + J\right]\frac{d}{dz} +{}\right.\\
+ \left[\frac{1}{2}\alpha JE\left(P + \frac{1}{2}\right) + \frac{1}{2}JE\right]\frac{d^2}{dz^2} + \left.\frac{1}{4}J^2E\frac{d^3}{dz^3}\right\}\mu +{}\\
+ \left\{\alpha\left(P + \frac{1}{2}\right)\frac{d}{dz} + J\frac{d^2}{dz^2} +{}\right.\\
+ \frac{1}{2}\alpha JE\left(P + \frac{1}{2}\right)\frac{d^3}{dz^3} + \left.\frac{1}{4}J^2E\frac{d^4}{dz^4}\right\}\varkappa = 0.
\end{aligned} \tag{8.27}$$

Здесь z — безразмерное время, определенное вторым из соотношений (8.9).

Динамика цепи регулирования зависит от характеристик датчика давления, передаточной функции усилителя и свойств регулятора и исполнительного привода. Мы не ставим здесь своей задачей подробно рассматривать конструкцию цепи обратной связи системы; поэтому определим динамику всей цепи регулирования с помощью операторного уравнения

$$F\left(\frac{d}{dz}\right)\varphi = \varkappa, \tag{8.28}$$

в котором F означает отношение двух многочленов, причем степень знаменателя выше степени числителя.

Уравнения (8.7), (8.27) и (8.28) образуют систему трех уравнений относительно трех переменных φ, μ и $\varkappa$. Так как эти уравнения принадлежат к классу уравнений с постоянными коэффициентами, целесообразно искать их решение в виде

$$\varphi = ae^{sz}; \qquad \mu = be^{sz}; \qquad \varkappa = ce^{sz}. \tag{8.29}$$

Подставив (8.29) в уравнения (8.7), (8.27) и (8.28), мы получим три однородных уравнения относительно параметров a, b и c:

$$a\left[s+(1-n)+ne^{-\delta s}\right]-be^{-\delta s}=0,$$

$$P\left[1+\alpha E\left(P+\frac{1}{2}\right)s+\frac{1}{2}JEs^2\right]+a\left\{\left[1+\alpha\left(P+\frac{1}{2}\right)\right]+\right.$$
$$+\left[\alpha E\left(P+\frac{1}{2}\right)+J\right]s+\left[\frac{1}{2}\alpha JE\left(P+\frac{1}{2}\right)+\frac{1}{2}JE\right]s^2+$$
$$\left.+\frac{1}{4}J^2Es^3\right\}b+s\left[\alpha\left(P+\frac{1}{2}\right)+Js+\frac{1}{2}\alpha JE\left(P+\frac{1}{2}\right)s^2+\right.$$
$$\left.+\frac{1}{4}J^2Es^3\right]c=0,$$

$$F(s)\,a-c=0.$$

Для того чтобы последняя система имела нетривиальное решение по a, b и c, определитель, составленный из коэффициентов при этих неизвестных, должен обращаться в нуль. Условие обращения в нуль этого определителя

записывается в виде

$$
\begin{gathered}
[s+(1-n)]\left\{\frac{1}{4}J^2Es^3+\frac{1}{2}JE\left[1+\alpha\left(P+\frac{1}{2}\right)\right]s^2+\right.\\
\left.+\left[\alpha E\left(P+\frac{1}{2}\right)+J\right]s+\left[1+\alpha\left(P+\frac{1}{2}\right)\right]\right\}+\\
+e^{-\delta s}\left[\frac{1}{4}nJ^2Es^3+\left\{\frac{1}{2}nJE\left[1+\alpha\left(P+\frac{1}{2}\right)\right]+\frac{1}{2}JEP\right\}s^2+\right.\\
+\left\{n\left[\alpha E\left(P+\frac{1}{2}\right)+J\right]+\alpha EP\left(P+\frac{1}{2}\right)\right\}s+\\
+\left\{n\left[n+\alpha\left(P+\frac{1}{2}\right)\right]+P\right\}+\\
\left.+sF(s)\left\{\frac{1}{4}J^2Es^3+\frac{1}{2}\alpha JE\left(P+\frac{1}{2}\right)s^2+Js+\alpha\left(P+\frac{1}{2}\right)\right\}\right]=0.
\end{gathered}
\tag{8.30}
$$

Равенство (8.30) представляет собой уравнение относительно показателя степени s, а функция $F(s)$ – передаточную функцию всей цепи обратной связи. Ответ на вопрос об устойчивости полной системы зависит от того, имеет ли уравнение (8.30) корни с положительной вещественной частью, или не имеет таких корней.

8.4. Неустойчивость системы в отсутствие цепи обратной связи. Характеристическое уравнение системы при разомкнутой цепи обратной связи можно простым образом получить с помощью основного уравнения (8.30), положив $F(s)=0$. Допустим, что многочлен, служащий коэффициентом при экспоненте $e^{-\delta s}$, не имеет нулей в правой полуплоскости комплексной переменной s, что обычно и имеет место. Тогда обе части уравнения (8.30) можно разделить на этот многочлен, не вводя при этом в результат полюсов, лежащих в правой полуплоскости переменной s.

Иными словами, при построении диаграммы Сатче мы вновь разбиваем функцию $G(s)$ таким образом, что

$$
G(s)=g_1(s)-g_2(s); \qquad g_1(s)=e^{-\delta s}.
$$

Кривая $g_1(s)$ попрежнему представляет собой окружность единичного радиуса. Выражение для функции $g_2(s)$ имеет

в данном случае более сложный вид:

$$g_2(s) = -\frac{g_2^{(1)}(s)}{g_2^{(2)}(s)},$$

где

$$g_2^{(1)}(s) = \left(\frac{s}{n} + \frac{1-n}{n}\right)\left\{\frac{1}{4}J^2Es^3 + \frac{1}{2}JE\left[1+\alpha\left(P+\frac{1}{2}\right)\right]s^2 + \right.$$
$$\left. + \left[\alpha E\left(P+\frac{1}{2}\right)+J\right]s + \left[1+\alpha\left(P+\frac{1}{2}\right)\right]\right\};$$

$$g_2^{(2)}(s) = \frac{1}{4}J^2Es^3 + \frac{1}{2}JE\left[1+\alpha\left(P+\frac{1}{2}\right)+\frac{P}{n}\right]s^2 +$$
$$+ \left[\alpha E\left(P+\frac{1}{2}\right)\left(1+\frac{P}{n}\right)+J\right]s + \left[1+\alpha\left(P+\frac{1}{2}\right)+\frac{P}{n}\right]. \tag{8.31}$$

При значениях s, стремящихся к $s=0$ вдоль мнимой оси, кривая $g_2(s)$ пересекает вещественную ось в точке, определяемой путем подстановки в (8.31) значения $s=0$, т. е. в точке

$$g_2(0) = -\frac{1-n}{n}\,\frac{1+\alpha\left(P+\frac{1}{2}\right)}{1+\alpha\left(P+\frac{1}{2}\right)+\frac{P}{n}}. \tag{8.32}$$

Так как параметры n, α и P положительны, то найденная абсолютная величина $g_2(0)$ в данном случае меньше абсолютной величины $g_2(0)$, определяемой с помощью соотношения (8.14), которое характеризует условие устойчивости двигателя. Таким образом, влияние системы подачи топлива выражается в смещении кривой $g_2(s)$ по направлению к единичному кругу $g_1(s)$ на диаграмме Сатче. Например, когда $n=\frac{1}{2}$, кривая $g_2(s)$ в точности касается единичного круга, соответствующего собственно двигателю, т. е. двигателю без подключенной к нему насосной (питающей) установки. При подключении этой установки кривая $g_2(s)$ уже пересекает окружность единичного радиуса, и при значениях временного запаздывания δ, превышающих некоторую конечную величину, система становится неустойчивой.

Таким образом, установка для подачи топлива всегда оказывает дестабилизирующее влияние. Это обстоятель-

ство становится еще более ясным, если обратиться к рассмотрению асимптоты кривой $g_2(s)$, которую мы получим с помощью соотношения (8.31) при чисто мнимых значениях s, больших по абсолютной величине. Мы будем иметь

$$g_2(i\omega) \approx -\left[\frac{i\omega}{n} + \left(\frac{1-n}{n} - \frac{2P}{Jn^2}\right) + \dots\right] \text{ при } |\omega| \gg 1. \quad (8.33)$$

Следовательно, при чисто мнимых значениях s, больших по абсолютной величине, кривая $g_2(s)$ асимптотически приближается к прямой, параллельной мнимой оси и расположенной от нее слева на расстоянии

$$\frac{1-n}{n} - \frac{2P}{Jn^2}.$$

Влияние подключения к двигателю установки для подачи топлива опять сводится к сдвигу кривой $g_2(s)$ по направлению к единичному кругу. Таким образом, очевидно, что для значений параметра n, близких к $1/2$ или больших $1/2$, невозможно сконструировать систему с безусловной устойчивостью. На диаграмме Сатче кривые $g_1(s)$ и $g_2(s)$ всегда будут пересекаться, если обратная связь отсутствует.

8.5. Устойчивость системы с обратной связью. Если многочлен

$$\begin{aligned} H(s) = \frac{1}{4} J^2 E s^3 + \left\{\frac{1}{2} JE\left[1 + \alpha\left(P + \frac{1}{2}\right)\right] + \frac{JEP}{2n}\right\} s^2 + \\ + \left[\alpha E\left(P + \frac{1}{2}\right) + \frac{\alpha EP}{n}\left(P + \frac{1}{2}\right)\right] s + \\ + \left[1 + \alpha\left(P + \frac{1}{2}\right) + \frac{P}{n}\right] + \\ + \frac{sF(s)}{n}\left[\frac{1}{4} J^2 E s^3 + \frac{1}{2}\alpha JE\left(P + \frac{1}{2}\right) s^2 + Js + \alpha\left(P + \frac{1}{2}\right)\right], \end{aligned} \quad (8.34)$$

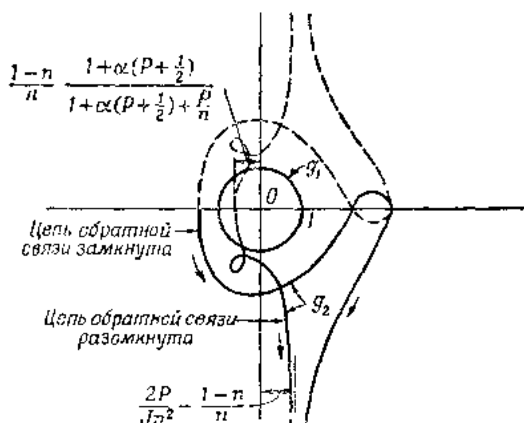
служащий коэффициентом при $e^{-\tau s}$ в левой части уравнения (8.30), не имеет ни нулей, ни полюсов в правой части полуплоскости комплексной переменной s, то с помощью диаграммы Сатче можно установить, имеет ли уравнение (8.30) нули в этой полуплоскости или не имеет. Для этого в качестве g_1 и g_2 выберем функции

$$g_1(s) = e^{-\tau s}$$

и

$$g_2(s) = -\left(\frac{s}{n} + \frac{1-n}{n}\right)\frac{1}{H(s)}\left\{\frac{1}{4}J^2Es^3 + \right.$$
$$+\frac{1}{2}JE\left[1+\alpha\left(P+\frac{1}{2}\right)\right]s^2 + \left[\alpha E\left(P+\frac{1}{2}\right)+J\right]s +$$
$$\left. + \left[1+\alpha\left(P+\frac{1}{2}\right)\right]\right\}. \qquad (8.35)$$

Когда s обегает контур, показанный на фиг. 31, кривая $g_1(s)$ попрежнему сводится к окружности единичного



Фиг. 57

радиуса. Следовательно, если при этом кривая $g_2(s)$ полностью лежит вне единичного круга, уравнение (8.30) не имеет корней в правой полуплоскости комплексной переменной s. Другими словами, система с обратной связью будет устойчива при любых значениях временнóго запаздывания при таком строении цепи обратной связи [характеризуемой передаточной функцией $F(s)$], при котором кривая $g_2(s)$ будет целиком расположена вне единичного круга (см. фиг. 57).

В качестве примера зададим следующие значения параметров:

$$n = \frac{1}{2}; \qquad P = \frac{3}{2}; \qquad J = 4; \qquad E = \frac{1}{4}; \qquad \alpha = 1,$$

Данному значению α отвечает центробежный насос в установке подачи жидкого топлива. Тогда без применения автоматического регулирования давления[1]) функция $g_2(s)$ имеет вид

$$g_2(s) = -\frac{1}{2}\,\frac{(2s+1)(2s^3+3s^2+9s+6)}{s^3+3s^2+6s+6}.$$

Для нашего исследования имеет первостепенное значение поведение $g_2(s)$ при чисто мнимых значениях $s = i\omega$, где ω — вещественная переменная. Тогда

$$g_2(i\omega) = -\frac{1}{2}\,\frac{(6-21\omega^2+4\omega^4)(6-3\omega^2)+\omega^2(21-8\omega^2)(6-\omega^2)}{(6-3\omega^2)^2+\omega^2(6-\omega^2)^2} -$$

$$-\frac{1}{2}\,i\omega\,\frac{(21-8\omega^2)(6-3\omega^2)-(6-21\omega^2+4\omega^4)(6-\omega^2)}{(6-3\omega^2)^2+\omega^2(6-\omega^2)^2}.$$

Часть этой кривой, соответствующая $\omega > 0$, нанесена на фиг. 58. Очевидно, что при достаточно больших значениях временно́го запаздывания система будет неустойчива. С другой стороны, если характер кривой $g_2(s)$ можно изменить путем введения цепи автоматического регулирования[2]) таким образом, чтобы $g_2(s)$ приняла, например, вид

$$g_2(s) = -2\,\frac{(s+2)(s+3)}{s+6},$$

то, как следует из рассмотрения фиг. 58, новая кривая g_2 целиком пройдет вне единичного круга g_1. Следовательно, в последнем случае система обладает абсолютной устойчивостью. Подставив выбранные значения параметров в соотношения (8.31) и (8.35), мы найдем требуемую передаточную функцию цепи автоматического регулирования:

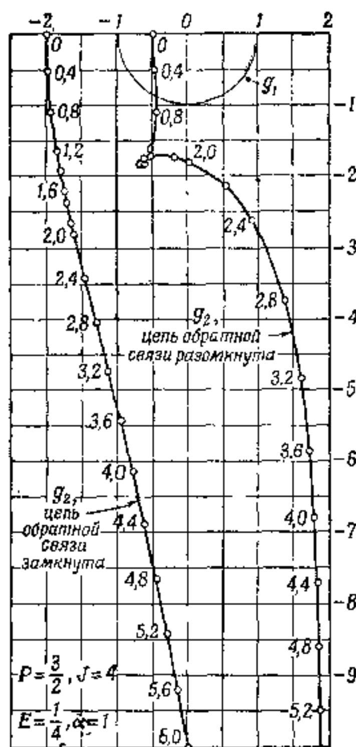
$$F(s) = -4{,}875\,\frac{(s+1{,}0528)(s^2+0{,}7164s+2{,}6301)}{s(s-2)(s+3)(s+0{,}5332)(s^2+0{,}4668s-3{,}1511)}.$$

Таким образом, эта цепь имеет характер интегрирующей цепи, рассмотренной в п. 3.3. Если при заданных характеристиках датчика (измерителя) давления и регулятора,

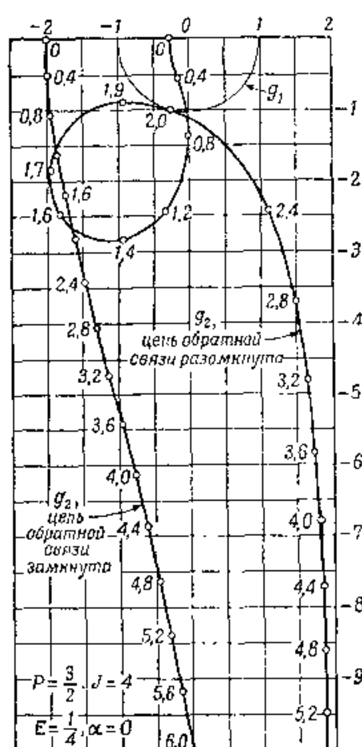
[1]) То есть при разомкнутой цепи обратной связи на фиг. 56. — *Прим. перев.*

[2]) То есть при замыкании цепи обратной связи на фиг. 56. — *Прим. перев.*

исполнительный привод которого управляет регулирующей емкостью, можно разработать такой усилитель, чтобы передаточная функция всей цепи автоматического регулирования в достаточной степени приближалась к только что



Фиг. 58



Фиг. 59

найденной передаточной функции, то имеется возможность стабилизировать процесс горения с помощью автоматического регулятора.

Во втором примере зададим значения

$$n=\frac{1}{2}; \qquad P=\frac{3}{2}; \qquad J=4; \qquad E=\frac{1}{4}; \qquad \alpha=0.$$

Так как $\alpha=0$, то подача топлива происходит под постоянным давлением p_0 даже при изменении расхода топлива,

что соответствует случаю простого нагнетателя. Тогда без применения автоматического регулятора

$$g_2(s) = -\frac{1}{2}\,\frac{(2s+1)(2s^3+s^2+8s+2)}{s^3+2s^2+4s+4}.$$

Если s принимает чисто мнимые значения, то

$$g_2(i\omega) = -\frac{1}{2}\,\frac{(4-2\omega^2)(2-17\omega^2+4\omega^4)+\omega^2(4-\omega^2)(12-4\omega^2)}{(4-2\omega^2)^2+\omega^2(4-\omega^2)^2} -$$
$$-\frac{1}{2}\,i\omega\,\frac{(4-2\omega^2)(12-4\omega^2)-(4-\omega^2)(2-17\omega^2+4\omega^4)}{(4-2\omega^2)^2+\omega^2(4-\omega^2)^2}.$$

Кривая g_2, соответствующая последнему уравнению, нанесена на фиг. 59. Очевидно, что без применения автоматического регулятора процесс горения при достаточно большом временнóм запаздывании будет неустойчивым. В самом деле, эта система обладает еще меньшим запасом устойчивости, чем система, рассмотренная в первом примере, в том смысле, что в ней неустойчивость наступит при еще меньшем запаздывании. Особый интерес представляет часть кривой g_2 в окрестности значения $\omega = 2$. Вблизи точки $\omega = 2$ эта кривая подходит так близко к единичному кругу g_1, что при соответствующем значении временнóго запаздывания окажется $g_1 \approx g_2$ и при $\omega \approx 2$ возникнут почти незатухающие колебания. Это критическое значение δ, очевидно, меньше его другого критического значения, соответствующего точке пересечения кривой g_2 с окружностью единичного радиуса при $\omega \approx 0{,}65$.

Для обеспечения абсолютной устойчивости кривую g_2 нужно сместить так, чтобы она не проходила больше через круг единичного радиуса, например сместить ее к «устойчивой» кривой, которую мы имели в первом примере. Для этого необходимо замкнуть систему с помощью цепи обратной связи с передаточной функцией

$$F(s) = -4{,}875\,\frac{(s+0{,}8126)(s^2-0{,}01337s+2{,}6506)}{s^2(s+2)(s+3)(s^2+4)}.$$

Следовательно, эта цепь должна иметь характер двойной интегрирующей цепи. Далее, ее передаточная функция $F(s)$ должна иметь в точках $\pm 2i$ два чисто мнимых полюса. Это требование к строению усилителя, не выполнимое в реальных физических устройствах, обусловлено характером системы подачи топлива и непосредственно вызвано

пренебрежением трением в питающем трубопроводе. В реальных же физических устройствах влияние трения в трубопроводе приводит к исчезновению этой пары мнимых полюсов искомой передаточной функции $F(s)$ и замене их двумя комплексными сопряженными полюсами[1]).

Следует подчеркнуть, что преимущество применения системы с обратной связью для стабилизации процесса горения в этой задаче состоит в ее большой гибкости в смысле возможности достижения абсолютной устойчивости (т. е. устойчивости при любых значениях параметра δ или τ). А так как мы не располагаем достоверными данными о величине временно́го запаздывания, то возможность достижения абсолютной устойчивости имеет большое практическое значение. Больше того, с помощью автоматического регулирования открывается возможность проектировать систему даже так, чтобы она оставалась устойчивой при всех ожидаемых изменениях параметра n. Из физических соображений следует, что n может принимать значения, лежащие между $^1/_2$ и 1. В худшем случае систему следует рассчитывать на значение $n \approx 1$. Тогда система будет устойчива при всех ожидаемых значениях n. Поэтому с помощью автоматического регулятора можно обеспечить устойчивость процесса горения и без точных данных о значениях параметров системы[2]).

8.6. Общий критерий устойчивости для систем с запаздывающим аргументом. В предыдущих рассуждениях о стабилизации посредством автоматического регулирова-

---

[1]) Конечно, с отрицательной вещественной частью, характеризующей затухание свободных колебаний системы за счет гидравлических потерь, обусловленных трением в трубопроводе.—*Прим. перев.*

[2]) Подчеркнем, что здесь речь идет о фиксированных, хотя и неизвестных нам в точности значениях параметров. Если же в процессе регулирования эти параметры могут претерпевать изменения хотя бы и в пределах тех значений, при которых обеспечена устойчивость системы в предположении постоянства параметров, то система будет системой с переменными параметрами и данный метод не пригоден для исследования ее устойчивости. Примеры неприменимости метода «замораживания» коэффициентов в системах с переменными параметрами приведены, например, в работе [46] и на стр. 46 сборника [45].—*Прим. перев.*

ния предполагалось, что многочлен $H(s)$, определяемый соотношением (8.34), не имеет ни нулей, ни полюсов в правой части плоскости комплексной переменной s. Однако это предположение выполняется не всегда. В общем случае нужно прежде всего узнать число нулей и полюсов функции $H(s)$ в этой полуплоскости. Для этого заметим, что сумма тех слагаемых многочлена $H(s)$, которые не содержат члена, пропорционального $F(s)$, обычно не имеет нулей в правой полуплоскости комплексной переменной s. Поэтому вместо изучения $H(s)$ можно изучать отношение $H(s)$ к многочлену, представляющему собой сумму упомянутых слагаемых. Это означает, что сумма числа нулей и числа полюсов функции $H(s)$ в правой полуплоскости комплексной переменной s равна сумме числа нулей и числа полюсов следующей функции:

$$H(s):\Big[\frac{1}{4}J^2Es^3+\Big\{\frac{1}{2}JE\Big[1+\alpha\Big(P+\frac{1}{2}\Big)\Big]+\frac{JEP}{2n}\Big\}s^2+ \\ +\Big[\alpha E\Big(P+\frac{1}{2}\Big)+\frac{\alpha EP}{n}\Big(P+\frac{1}{2}\Big)\Big]s+\Big[1+\alpha\Big(P+\frac{1}{2}\Big)+\frac{P}{n}\Big]\Big]= \\ =1+L(s), \quad (8.36)$$

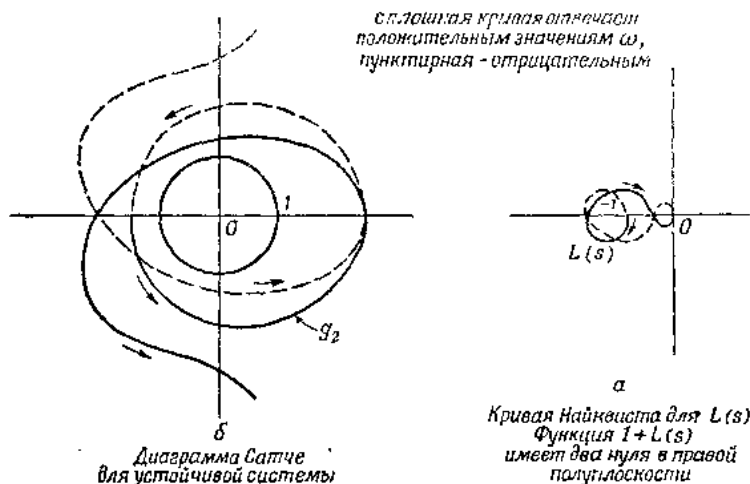
где

$$L(s)=\frac{1}{n}sF(s)\Big[\frac{1}{4}J^2Es^3+ \\ +\frac{1}{2}\alpha JE\Big(P+\frac{1}{2}\Big)s^2+Js+\alpha\Big(P+\frac{1}{2}\Big)\Big]: \\ \Big[\frac{1}{4}J^2Es^3+\Big\{\frac{1}{2}JE\Big[1+\alpha\Big(P+\frac{1}{2}\Big)\Big]+\frac{JEP}{2n}\Big\}s^2+ \\ +\Big[\alpha E\Big(P+\frac{1}{2}\Big)+\frac{\alpha EP}{n}\Big(P+\frac{1}{2}\Big)\Big]s+ \\ +\Big[1+\alpha\Big(P+\frac{1}{2}\Big)+\frac{P}{n}\Big]\Big]. \quad (8.37)$$

Согласно критерию Найквиста, число нулей и полюсов функции $1+L(s)$ в правой полуплоскости s можно найти путем построения кривой Найквиста для функции $1+L(s)$ в предположении, что s описывает кривую, изображенную на фиг. 31. В самом деле, если функция $1+L(s)$ или $H(s)$ имеет r нулей и q полюсов в правой полуплоскости s, то при обходе точкой s полуокружности

вектор $L(s)$ совершит $r-q$ оборотов по часовой стрелке вокруг точки -1. Отсюда следует, что кривая Найквиста для функции $L(s)$ доставляет все необходимые сведения относительно функции $H(s)$.

При делении на $H(s)$ функции, стоящей в левой части равенства (8.30), с целью получения функций $g_1(s)$ и $g_2(s)$ [см. (8.35)], в правой части плоскости комплексной переменной s появляется q нулей и r полюсов. q полюсов



Фиг. 60

функции $L(s)$ появляются через посредство функции $F(s)$, так как многочлен, находящийся в знаменателе правой части равенства (8.37), не имеет нулей в правой полуплоскости комплексной переменной s. Поэтому выражение, стоящее в левой части равенства (8.30), также имеет q полюсов в правой полуплоскости комплексной переменной s. Следовательно, чтобы это выражение не имело нулей в правой полуплоскости, $g_2(s)$ должно совершить $-q+(q-r)=-r$ оборотов по часовой стрелке (r оборотов против часовой стрелки) вокруг единичного круга. Для того чтобы устойчивость была абсолютной, т. е. сохранялась при любых значениях временного запаздывания, кривая $g_2(s)$ не должна никогда пересекать окружности единичного радиуса. Поэтому *общий критерий абсолют-*

ной устойчивости выражается следующими двумя условиями: 1) кривая $g_2(s)$ должна целиком лежать вне единичного круга; 2) кривая $g_2(s)$ должна r раз оборачиваться вокруг единичного круга против часовой стрелки, когда s обходит обычный контур в правой полуплоскости комплексной переменной s. Этот критерий устойчивости основан на применении диаграммы Сатче. Число r определяется с помощью кривой Найквиста для функции $L(s)$, заданной соотношением (8.37). Таким образом, исследование устойчивости в общем случае системы с запаздывающим аргументом требует привлечения как диаграммы Сатче, так и диаграммы Найквиста (фиг. 60).

Очевидно, что сформулированный здесь критерий устойчивости, основанный на совместном использовании диаграмм Сатче и Найквиста, применим к любой системе, содержащей временное запаздывание τ. Вопрос об устойчивости таких систем всегда сводится к исследованию задачи о том, имеет или не имеет уравнение

$$M(s) = 0$$

корни с положительной вещественной частью; при этом функция $M(s)$ содержит члены, имеющие множителем экспоненту $e^{-\tau s}$. Как следует из проведенного рассуждения, существо метода состоит в том, что функция $M(s)$ делится на коэффициент при $e^{-\tau s}$ в выражении, определяющем $M(s)$, так что

$$\frac{M(s)}{1 + L(s)} = G(s) = g_1(s) - g_2(s)$$

и

$$g_1(s) = e^{-\tau s}.$$

Теперь диаграмма Сатче определяется совокупностью кривых, описываемых функциями $g_1(s)$ и $g_2(s)$, когда переменная s описывает полуокружность в правой полуплоскости, изображенную на фиг. 31, причем кривая, описываемая функцией $g_1(s)$, сводится к окружности единичного радиуса. Чтобы узнать, не привело ли деление функции $M(s)$ на функцию $1 + L(s)$ к появлению корней с положительными вещественными частями, влияющими на построение диаграммы Сатче, мы должны построить кривую Найквиста для функции $L(s)$. Число корней уравнения

$M(s)=0$, имеющих положительную вещественную часть, определяется с помощью теоремы Коши.

Передаточная функция цепи обратной связи входит в выражение для $g_2(s)$; при этом частью этой передаточной функции, которая зависит от строения усилителя, проектировщик может располагать по своему усмотрению. Функция $g_2(s)$ может также содержать трансцендентные функции от s, отражающие влияние других частей системы. Так как передаточная функция усилителя в цепи обратной связи представляет собой в общем случае отношение двух многочленов по степеням s, достижение полной компенсации дестабилизирующего влияния, обусловленного трансцендентной функцией, связано с большими затруднениями. Однако на диаграмме Сатче критическая часть кривой, описываемой функцией $g_2(s)$, проходит в области, в которой она оказывается ближе всего к единичному кругу $g_1(s)$. Но значения $g_2(s)$, близкие к единичному кругу, в общем случае соответствуют малым по абсолютной величине значениям s. Отсюда следует, что в области критической части кривой $g_2(s)$ трансцендентную функцию можно разложить в ряд по степеням s. Благодаря этому при проектировании усилителя в цепи обратной связи можно исходить из приближенного выражения ряда при замене его суммой нескольких членов[1]. Таким образом, усилитель компенсирует дестабилизирующее действие основной части системы в критической области. Само собой разумеется, что в конечном счете необходимо проверить качество системы с помощью изложенного выше критерия устойчивости, учитывая характеристики построенного таким образом усилителя. За дальнейшими подробностями мы отсылаем к работе Ф. Э. Марбла[2], предложившего указанный здесь метод.

---

[1] По поводу этой методики см. предостережение Я. З. Цыпкина [29]. — *Прим. перев.*

[2] Marble F. E., Cox D. W., Journ. Amer. Rock. Soc., **24**, 75—81 (1953).

Глава IX

ЛИНЕЙНЫЕ СИСТЕМЫ, НАХОДЯЩИЕСЯ ПОД ДЕЙСТВИЕМ СТАЦИОНАРНЫХ СЛУЧАЙНЫХ ВХОДНЫХ ФУНКЦИЙ

В предыдущих главах входные воздействия рассматривались как вполне определенные функции времени t. Однако существует много технических задач, связанных с исследованием линейных систем с постоянными коэффициентами, в которых входные воздействия невозможно выразить в столь определенной форме. Такая задача возникает, например, при изучении деформации крыла самолета и напряжений в материале этого крыла в условиях обтекания его турбулентным потоком воздуха. Здесь за вход можно считать параметры потока воздуха, меняющиеся со временем. Но турбулентное течение воздуха нельзя описать с помощью определенной функции времени; параметры такого течения приходится рассматривать как случайные функции времени, подчиненные некоторым статистическим закономерностям. Очевидно, что тогда и выход системы (в данном случае напряжения в материале крыла) также должен быть случайной функцией и поддается описанию только с помощью статистических методов. Поэтому первая цель настоящей главы состоит в разыскании надлежащего метода определения статистических свойств выхода по заданным статистическим свойствам входа. Такой метод является очевидным обобщением ранних исследований П. Ланжевена о броуновском движении.

Другим примером случайного входного воздействия служат так называемые шумы, налагающиеся на управляющий сигнал и возникающие за счет возмущений и флуктуаций, устранить которые из реальной системы проектировщик не в силах. Проблема шумов составляет предмет весьма обширных исследований в области техники связи. Главная задача таких исследований заключается в том, чтобы свести к минимуму влияние неустранимых

шумов без подавления полезной информации, содержащейся в сигнале; этой специфической задачей фильтрации шумов мы займемся в гл. XVI. Задача же, изучаемая в настоящей главе, имеет несколько отличный характер. В рассматриваемой здесь задаче весь выход системы сводится к случайной функции. Наша задача в области разработки систем, в особенности систем с обратной связью, заключается в том, чтобы получить на выходе сигнал, обладающий желаемыми статистическими характеристиками, коль скоро задан входной сигнал. Мы увидим, что метод передаточных функций, развитый в предыдущих главах, остается полезным и при решении настоящей задачи.

9.1. Статистическое описание случайных функций. Рассмотрим систему, вырабатывающую случайную функцию $y_1(t)$. Чтобы установить способ статистического описания такой случайной функции, нам придется рассмотреть большое число систем, тождественных только что указанной. Мы будем называть совокупность систем множеством систем. Случайные функции, вырабатываемые элементами, принадлежащими к этому множеству, обозначим $y_1(t)$, $y_2(t)$, $y_3(t)$, ... Случайный характер функции проявляется в том, что хотя все эти системы тождественны, однако значение функции, вырабатываемой любым элементом рассматриваемого множества в любой определенный момент времени t, в общем случае отличается от значения, выработанного другим элементом в тот же самый момент. Но мы можем поставить вопрос о том, какая часть полного числа систем вырабатывает значение, лежащее внутри заданного промежутка от y до $y+dy$. Эта часть зависит от величин y и t и при малых значениях dy будет пропорциональна dy и равна вероятности нахождения y между значениями y и $y+dy$ в момент t. Обозначим ее через $W_1(y, t)\,dy$. Функцию $W_1(y, t)$ назовем первой функцией распределения вероятности. Далее, мы можем рассмотреть все пары значений y, соответствующие двум моментам t_1 и t_2. Часть полного числа пар, которым соответствуют значения y, заключенные внутри промежутка от y_1 до y_1+dy_1 в момент t_1 и внутри промежутка от y_2 до y_2+dy_2 в момент t_2, обозначим через $W_2(y_1, t_1; y_2, t_2)\,dy_1\,dy_2$. Функцию $W_2(y_1, t_1; y_2, t_2)$ назо-

вем второй функцией распределения вероятности. Подобным же образом строятся и функции распределения вероятностей более высоких порядков.

Можно возразить, что принятое построение статистического описания случайной функции, основанное на одновременном наблюдении за работой весьма большого числа тождественных систем, наталкивается на трудности, связанные с практическим осуществлением такого наблюдения. Однако если случайная функция является стационарной случайной функцией в том смысле, что все ее статистические свойства не зависят от времени, то нет надобности прибегать к введению многих систем, а все необходимые наблюдения можно провести над одной системой в течение достаточно длительного промежутка времени. Затем запись результатов наблюдения можно разрезать на отдельные куски длины θ, где величина θ велика по сравнению с характеристическим временем функции. Следовательно, каждый такой кусок записи доставляет одну и ту же статистическую информацию о поведении системы, так как при статистическом характере изменения переменной расположение начала на оси времени не имеет значения. Таким образом, различные куски записи можно рассматривать как множество наблюдений над тождественными системами, причем можно определить различные функции распределения вероятности. Кроме того, эти функции распределения в рассматриваемых условиях несколько упрощаются: в самом деле, величина W_1 не будет зависеть от времени, а величина W_2 будет зависеть лишь от интервала времени $\tau = t_2 - t_1$. Отсюда следует, что в случае стационарных случайных функций, $W_1(y)\,dy$ определяет вероятность нахождения y между y и $y+dy$; $W_2(y_1, y_2; \tau)\,dy_1 dy_2$ определяет вероятность нахождения пары величин между y_1 и y_1+dy_1 и между y_2 и y_2+dy_2, причем две величины разделены интервалом времени τ. Так как в инженерных задачах случайные функции можно очень часто принимать за случайные стационарные функции, то в дальнейшем мы ограничимся рассмотрением именно таких случайных функций.

Следует подчеркнуть, что функции распределения вероятностей заключают в себе все сведения о статистических свойствах случайной функции. Иначе можно сказать,

что функции распределения вероятностей «определяют» случайную функцию. Разумеется, что эти функции W_n не могут задаваться произвольно, а должны удовлетворять следующим условиям:

а) $W_n > 0$, так как отрицательные значения вероятности не существуют.

б) Функция W_n симметрична по отношению к ее аргументам y_i, т. е.

$$W_2(y_1, y_2; \tau) = W_2(y_2, y_1; \tau). \tag{9.1}$$

Это свойство вытекает из смысла W_2 как совместной функции распределения вероятностей.

в) Высшие функции распределения определенным образом связаны с низшими функциями:

$$\int_{-\infty}^{\infty} W_2(y_1, y_2; \tau)\, dy_2 = W_1(y_1) = W_1(y), \tag{9.2}$$

причем интегрирование распространяется по всем возможным значениям y_2. Заметим, что интегрированием по y_2 мы исключаем τ. Кроме того,

$$\int_{-\infty}^{\infty} W_1(y)\, dy = 1. \tag{9.3}$$

Последнее равенство выражает тот простой факт, что вероятность по всем значениям y сводится к достоверности.

9.2. Среднее значение. Исходя из первой функции $W_1(y)$ распределения вероятностей можно найти среднее значение $\bar{y}$ величины y:

$$\bar{y} = \int_{-\infty}^{\infty} y W_1(y)\, dy. \tag{9.4}$$

Так как мы условились рассматривать только стационарные случайные функции, средние значения можно также вычислять как средние по времени значения функции

$y(t)$, т. е.

$$\bar{y} = \lim_{\Theta\to\infty} \frac{1}{\Theta} \int_{-\frac{\Theta}{2}}^{\frac{\Theta}{2}} y(t)\,dt. \tag{9.5}$$

Равенство среднего значения (9.4) по множеству[1]) и среднего значения (9.5) по времени является характеристическим свойством стационарных случайных функций, которым мы будем неоднократно пользоваться в дальнейших вычислениях.

Соотношение (9.4) можно обобщить на случай произвольных степеней y. Таким образом

$$m_n = \overline{y^n} = \int_{-\infty}^{\infty} y^n W_1(y)\,dy. \tag{9.6}$$

Величина m_n называется моментом n-й степени первой функции распределения вероятностей. С помощью моментов первой и второй степеней мы можем подсчитать величину, называемую флуктуацией или средним отклонением σ:

$$\sigma^2 = \overline{(y-\bar{y})^2} = \int_{-\infty}^{\infty} (y-\bar{y})^2 W_1(y)\,dy =$$

$$= \int_{-\infty}^{\infty} [y^2 - 2y\bar{y} + (\bar{y})^2]\,W_1(y)\,dy = \overline{y^2} - (\bar{y})^2. \tag{9.7}$$

Таким образом, σ служит мерой «ширины» функции $W_1(y)$ распределения вероятностей около среднего значения $\bar{y}$. Подобным же образом момент третьей степени служит мерой несимметричности распределения вероятностей. С помощью моментов все более высоких степеней можно получить дальнейшие данные о функции $W_1(y)$. В некоторых случаях распределение вероятностей определяется

[1]) «Среднее по множеству» в русской литературе часто называется «математическим ожиданием».—*Прим. перев.*

с помощью моментов единственным образом. Пусть, например,

$$\left.\begin{aligned} m_{2k+1} &= 0, \qquad k = 0,\ 1,\ 2,\ \ldots, \\ m_{2k} &= 1\cdot 3\cdot 5\ldots(2k-1)\,\sigma^{2k}. \end{aligned}\right\} \tag{9.8}$$

В этом случае первая функция $W_1(y)$ распределения вероятностей сводится к хорошо известному гауссову закону распределения (или нормальному распределению)

$$W_1(y) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{y^2}{2\sigma^2}}. \tag{9.9}$$

В некоторых случаях оказывается удобным выбрать начало координат y таким образом, чтобы величина $\overline{y}$ обращалась в нуль, т. е. выбрать начало в точке, соответствующей среднему значению функции y. При выполнении этого условия мы говорим, что распределение вероятностей нормировано. В этом случае квадрат среднего отклонения представляет собой просто момент $\overline{y^2}$ второго порядка, что обнаруживается с помощью равенства (9.7).

Наиболее важным средним значением, образованным с помощью второй функции $W_2(y_1,\ y_2;\ \tau)$ распределения вероятностей, является функция $R(\tau)$ корреляции:

$$\begin{aligned} R(\tau) = \overline{y_1 y_2} = \overline{y(t)\,y(t+\tau)} = \\ = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y_1 y_2 W_2(y_1,\ y_2;\ \tau)\,dy_1\,dy_2. \end{aligned} \tag{9.10}$$

Для стационарных случайных функций последнее соотношение получается, конечно, также и с помощью осреднения по времени:

$$R(\tau) = \lim_{\Theta\to\infty}\frac{1}{\Theta}\int_{-\Theta/2}^{\Theta/2} y(t)\,y(t+\tau)\,dt. \tag{9.11}$$

Таким образом, функция $R(\tau)$ определяет степень взаимной связи двух значений функции y, измеренных в два различных момента времени. Следует ожидать, что с ростом интервала времени τ эта взаимная связь, или «память»,

должна ослабевать и, в конце концов, при достаточно больших значениях τ значения $y(t)$ и $y(t+\tau)$ перестанут зависеть одно от другого. Тогда, как известно из теории вероятностей, вторая функция распределения вероятностей будет равна произведению функций $W_1(y_1)$ и $W_1(y_2)$. Таким образом, при больших значениях τ

$$R(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y_1 y_2 W_1(y_1) W_1(y_2)\, dy_1\, dy_2 = (\overline{y})^2. \quad (9.12)$$

При $\tau = 0$ из равенства (9.11) следует, что

$$R(0) = \overline{y^2}. \quad (9.13)$$

Так как вычисление функции $R(\tau)$ можно производить при любом положении начала на оси времени, то имеют место соотношения

$$R(\tau) = \overline{y(t)\, y(t+\tau)} = \overline{y(t-\tau)\, y(t)}.$$

Дифференцируя $R(\tau)$ по времени и полагая затем $\tau = 0$, получаем

$$R'(0) = \overline{y(t)\, y'(t)} = -\overline{y(t)\, y'(t)}.$$

Поэтому

$$R'(0) = \overline{y(t)\, y'(t)} = 0. \quad (9.14)$$

В последних выражениях штрихи означают дифференцирование по времени. Таким образом, коэффициент корреляции случайной функции и ее производной, рассматриваемых в один и тот же момент времени, равен нулю. Это означает, что угловой коэффициент касательной к кривой, характеризующей течение функции y, может быть при любом значении y с одинаковой вероятностью положительным или отрицательным.

Дифференцируя $R(\tau)$ по τ два раза и полагая затем $\tau = 0$, находим, что

$$R''(0) = \overline{y(t)\, y''(t)} = -\overline{(y')^2}. \quad (9.15)$$

Последнее соотношение позволяет вычислять среднее квадратичное значение производной от y с помощью функции корреляции. Подобным же образом найдем среднее квадра-

тичное значение и второй производной от y:

$$R''''(0) = \overline{(y'')^2}. \qquad (9.16)$$

9.3. Спектральная плотность. Для дальнейших приложений теории случайных функций, которые мы имеем в виду, особо важную роль играет понятие о спектре случайной функции. Пусть мы наблюдаем функцию $y(t)$ в течение длительного промежутка времени Θ. В предположении о том, что вне этого промежутка $y(t)$ обращается в нуль, функцию $y(t)$ можно представить с помощью интеграла Фурье[1])

$$y(t) = \int_{-\infty}^{\infty} A(\omega)\, e^{i\omega t}\, d\omega, \qquad (9.17)$$

где $A(\omega)$ — амплитуда[2]) колебания частоты ω. Эту амплитуду можно подсчитать, зная $y(t)$, с помощью формулы обращения:

$$A(\omega) = \frac{1}{2\pi} \int_{-\Theta/2}^{\Theta/2} y(t)\, e^{-i\omega t}\, dt. \qquad (9.18)$$

Обозначим через $A^*(\omega)$ величину, комплексно сопряженную с $A(\omega)$; тогда, учитывая, что функция $y(t)$ является вещественной, из соотношения (9.18) выводим, что

$$A^*(\omega) = A(-\omega). \qquad (9.19)$$

Теперь мы имеем возможность выразить среднее квадратичное значение, т. е. величину $\overline{y^2}$, через посредство $A(\omega)$:

$$\overline{y^2} = \lim_{\Theta\to\infty} \frac{1}{\Theta} \int_{-\Theta/2}^{\Theta/2} y^2(t)\, dt =$$

$$= \lim_{\Theta\to\infty} \frac{1}{\Theta} \int_{-\Theta/2}^{\Theta/2} dt \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d\omega\, d\omega'\, A(\omega)\, A(\omega')\, e^{i(\omega+\omega')t}.$$

---

[1]) См., например, Уиттекер и Ватсон, Курс современного анализа, М.—Л., 1933.

[2]) Эта «амплитуда», вообще говоря, является комплексным числом. *Прим. ред.*

В результате замены переменных $\omega'' = -\omega'$ мы придем к выражению

$$\overline{y^2} = \lim_{\Theta \to \infty} \frac{1}{\Theta} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\omega\, d\omega''\, A(\omega)\, A^*(\omega'') \int_{-\Theta/2}^{\Theta/2} e^{i(\omega-\omega'')t}\, dt =$$

$$= \lim_{\Theta \to \infty} \frac{2}{\Theta} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A(\omega)\, A^*(\omega'') \frac{\sin\left[\frac{1}{2}(\omega-\omega'')\Theta\right]}{\omega-\omega''}\, d\omega\, d\omega''.$$

Введя теперь новую переменную ξ посредством соотношения

$$\xi = \frac{\Theta}{2}(\omega - \omega''),$$

получим, что

$$\omega'' = \omega - \frac{2\xi}{\Theta},$$

и, следовательно,

$$\overline{y^2} = \lim_{\Theta \to \infty} \frac{2}{\Theta} \int_{-\infty}^{\infty} A(\omega)\, d\omega \int_{-\infty}^{\infty} A^*\left(\omega - \frac{2\xi}{\Theta}\right) \frac{\sin \xi}{\xi}\, d\xi =$$

$$= \left[\lim_{\Theta \to \infty} \frac{1}{\Theta} \int_{0}^{\infty} |A(\omega)|^2\, d\omega\right] 4 \int_{-\infty}^{\infty} \frac{\sin \xi}{\xi}\, d\xi =$$

$$= 4\pi \lim_{\Theta \to \infty} \frac{1}{\Theta} \int_{0}^{\infty} |A(\omega)|^2\, d\omega.$$

Поэтому, если мы обозначим через $\Phi(\omega)$ функцию

$$\Phi(\omega) = \lim_{\Theta \to \infty} \frac{4\pi}{\Theta} |A(\omega)|^2, \tag{9.20}$$

то придем к следующему выражению среднеквадратичного значения:

$$\overline{y^2} = \int_{0}^{\infty} \Phi(\omega)\, d\omega. \tag{9.21}$$

Функция $\Phi(\omega)$ является, очевидно, вещественной функцией и носит название спектральной плотности случайной

функции. Соотношения (9.20) и (9.21) позволяют вычислять среднее значение $\overline{y^2}$ по изображению Фурье $A(\omega)$. Это свойство выражает теорему Парсеваля.

Рассмотрим далее корреляционную функцию $R(\tau)$. Сочетая соотношения (9.11) и (9.17), получаем

$$R(\tau)=\lim_{\Theta\to\infty}\frac{1}{\Theta}\int_{-\Theta/2}^{\Theta/2} y(t)\,y(t+\tau)\,dt=$$

$$=\lim_{\Theta\to\infty}\frac{1}{\Theta}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} A(\omega)\,A(\omega')\,e^{i\omega t}\,d\omega\,d\omega'\int_{-\Theta/2}^{\Theta/2} e^{i(\omega+\omega')t}\,dt.$$

Пользуясь рассуждением, аналогичным только что проведенному, мы находим, что

$$R(\tau)=\int_{0}^{\infty}\Phi(\omega)\cos\omega\tau\,d\omega. \tag{9.22}$$

Если в равенствах (9.21) и (9.22) положить $\tau=0$, то мы придем к равенству (9.13). Дифференцируя обе части (9.22) по τ и полагая затем $\tau=0$, приходим к (9.15). В силу теоремы обращения для преобразования Фурье[1]), справедливо соотношение

$$\Phi(\omega)=\frac{2}{\pi}\int_{0}^{\infty}R(\tau)\cos\omega\tau\,d\tau. \tag{9.23}$$

Соотношения (9.22) и (9.23) позволяют находить или функцию корреляции, или спектральную плотность, коль скоро задана одна из этих зависимостей; упомянутые равенства называются уравнениями Винера — Хинчина.

В графике функции спектральной плотности $\Phi(\omega)$ могут наблюдаться пики типа δ-функции Дирака[2]). Это обязательно произойдет, когда $\overline{y}$ отлично от нуля, или если пользоваться терминологией, принятой в электротехнике, когда y содержит составляющую постоянного

---

[1]) То есть в силу формулы типа (9.18), связывающей изображение Фурье с оригиналом. — *Прим. ред.*

[2]) О свойствах этой функции см. [48]. — *Прим. перев.*

тока. В этом случае

$$\Phi(\omega) = 2(\bar{y})^2 \delta(\omega) + \Phi_1(\omega), \tag{9.24}$$

где $\delta(x)$ определяется соотношениями

$$\left.\begin{aligned} &\delta(-x) = \delta(x) = 0 \qquad \text{при } x \neq 0, \\ &\delta(x) \to \infty \qquad \text{при } x = 0, \\ &\text{причем} \\ &\int_{-\infty}^{\infty} \delta(x)\,dx = 1 \quad \text{и} \quad \int_{0}^{\infty} \delta(x)\,dx = \frac{1}{2}. \end{aligned}\right\} \tag{9.25}$$

В случае чистого шума пик спектральной плотности, приходящийся на значение $\omega = 0$ и соответствующий постоянному току, обычно является единственным и, таким образом, функция $\Phi_1(\omega)$ будет регулярной и будет характеризовать подлинно непрерывный спектр. Но возможны также и такие случаи, когда на шум налагается несколько синусоидальных колебаний. Тогда спектральная плотность будет иметь дополнительные пики, образующиеся в изолированных точках, соответствующих частотам этих колебаний.

9.4. Примеры спектральной плотности. Рассмотрим несколько примеров вычисления спектральной плотности по функции корреляции. Если функция корреляции задана посредством кривой Гаусса

$$R(\tau) = R(0)\,e^{-\alpha^2\tau^2}, \tag{9.26}$$

то, в силу равенства (9.23), ей отвечает спектральная плотность

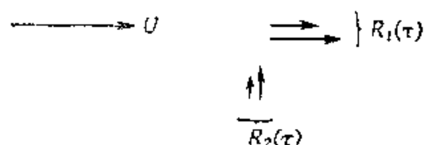
$$\Phi(\omega) = \frac{2}{\pi} R(0) \int_{0}^{\infty} \cos(\omega\tau)\, e^{-\alpha^2\tau^2}\, d\tau = \Phi(0)\, e^{-\frac{\omega^2}{4\alpha^2}}, \tag{9.27}$$

где

$$\Phi(0) = \frac{R(0)}{\alpha\sqrt{\pi}}.$$

Интересно отметить, что при $\alpha \to \infty$ функция корреляции обращается в нуль при всех конечных значениях τ, а $R(0) \to \infty$ таким образом, что $R(\tau)$ становится дельта-функцией. Это означает, что последующие значения y

уже не связаны никаким соотношением корреляции и что такая случайная функция является «наиболее хаотичной» из всех. При $\alpha \to \infty$ спектральная плотность становится постоянной, не зависящей от частоты. Эта наиболее хаотическая случайная функция называется белым шумом и часто оказывается пригодной для описания случайных изменений, происходящих в физических устройствах.



Фиг. 61

Другой пример связан с малой изотропной турбулентностью в потоке с постоянной скоростью. Как показали Карман и Хоуарт[1]), это течение характеризуется двумя основными функциями корреляции второго порядка $R_1(\tau)$ и $R_2(\tau)$: функция $R_1(\tau)$ представляет собой функцию корреляции составляющей пульсирующей скорости в направлении, параллельном среднему течению, для одной и той же точки пространства, но для моментов времени, отделенных промежутком в τ секунд (фиг. 61). Функция $R_2(\tau)$ представляет собой функцию корреляции составляющей пульсирующей скорости в направлении, перпендикулярном к среднему течению. Обозначая через U среднюю скорость течения, а через L — характерный размер турбулентности, мы можем выписать следующие приближенные выражения для этих функций корреляции:

$$R_1(\tau) = R_1(0)\, e^{-\frac{\tau U}{L}}, \tag{9.28}$$

$$R_2(\tau) = R_2(0)\, e^{-\frac{\tau U}{L}} \left(1 - \frac{1}{2}\frac{\tau U}{L}\right). \tag{9.29}$$

Исходя из соотношения (9.23), составим выражения для спектральных плотностей $\Phi_1(\omega)$ и $\Phi_2(\omega)$ флуктуаций составляющих скорости соответственно параллельной и пер-

[1]) Von Kármán and Howarth, Proc. Roy. Soc. (A), **164**, 192 (1938).

пендикулярной средней скорости:

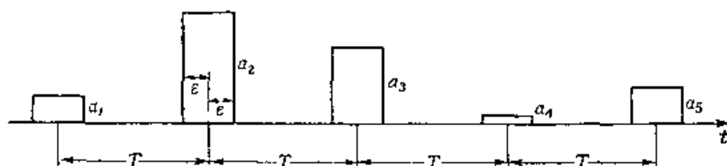
$$\Phi_1(\omega) = \Phi_1(0)\frac{1}{1+\left(\frac{\omega L}{U}\right)^2}, \qquad (9.30)$$

$$\Phi_2(\omega) = \Phi_2(0)\frac{1+3\left(\frac{\omega L}{U}\right)^2}{\left[1+\left(\frac{\omega L}{U}\right)^2\right]^2}, \qquad (9.31)$$

где $\Phi_1(0)$ и $\Phi_2(0)$ суть значения соответственных спектральных плотностей при $\omega = 0$. Эти значения связаны с величинами $R_1(0)$ и $R_2(0)$ посредством соотношений

$$\Phi_1(0) = \frac{2}{\pi}\frac{L}{U}R_1(0); \qquad \Phi_2(0) = \frac{1}{\pi}\frac{L}{U}R_2(0). \qquad (9.32)$$

9.5. Непосредственное вычисление спектральной плотности. Конечно, нет необходимости вычислять спектральную плотность по функции корреляции. Иногда возможно



Фиг. 62

находить эту плотность непосредственно по заданным свойствам самой случайной функции $y(t)$. Рассмотрим, например, функцию $y(t)$, определенную следующим образом: $y(t)$ представляет собой последовательность импульсов одинаковой формы и одинаковой периодичности, высоты которых меняются согласно некоторому вероятностному распределению. При этом высоты двух последовательных импульсов не связаны какой-либо корреляционной зависимостью. Такая функция изображена на фиг. 62 применительно к случаю прямоугольных импульсов. Если $\eta(t)$ означает одинарный импульс единичной высоты, то

$$y(t) = \sum_k a_k \eta(t-kT), \qquad (9.33)$$

где T — расстояние между каждыми двумя последовательными импульсами, а a_k — высота k-го импульса. Первый шаг при вычислении спектральной плотности заключается в нахождении спектра $A(\omega)$ Фурье по формуле (9.18). Положим $\theta = 2NT$, тогда

$$A(\omega) = \frac{1}{2\pi} \int_{-NT}^{NT} y(t)\, e^{-i\omega t}\, dt = \frac{1}{2\pi} \int_{-NT}^{NT} \sum_k a_k \eta(t - kT)\, e^{i\omega t}\, dt =$$

$$= \sum_{-N}^{N} a_k e^{-i\omega kT} \frac{1}{2\pi} \int_{-\infty}^{\infty} \eta(\xi)\, e^{-i\omega\xi}\, d\xi = \alpha(\omega) \sum_{-N}^{N} a_k e^{-i\omega kT},$$

где $\alpha(\omega)$ — спектр Фурье для единичного импульса,

$$\alpha(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \eta(\xi) e^{-i\omega\xi}\, d\xi. \tag{9.34}$$

Для прямоугольного импульса ширины 2ε и единичной высоты

$$\alpha(\omega) = \frac{1}{2\pi} \int_{-\varepsilon}^{\varepsilon} e^{-i\omega\xi}\, d\xi = \frac{1}{\pi} \frac{\sin \omega\varepsilon}{\omega}. \tag{9.35}$$

В силу равенств (9.19) и (9.20), спектральная плотность определяется соотношением

$$\Phi(\omega) = \frac{4\pi}{T} |\alpha(\omega)|^2 \lim_{N \to \infty} \frac{1}{2N} \left[\sum_{-N}^{N} \sum_{-N}^{N} a_k a_l e^{-i\omega(k-l)T} \right]. \tag{9.36}$$

Для выполнения предельного перехода в правой части (9.36) целесообразно предварительно перейти к среднему по множеству от всей этой правой части. Так как спектральная плотность $\Phi(\omega)$ одинакова для каждого элемента рассматриваемого множества, то левая часть равенства (9.36) при осреднении по множеству не меняется. Что касается правой части, то ее вид при выполнении такого осреднения упростится. Обозначим через $\overline{a}$ среднее значение a_k и a_l и через $\overline{a^2}$ — среднее значение квадратов a_k и a_l. Тогда

$$a_k a_l = (a_k - \overline{a})(a_l - \overline{a}) + \overline{a}\,[(a_k - \overline{a}) + (a_l - \overline{a})] + (\overline{a})^2.$$

Подставим это выражение в правую часть равенства (9.36) и затем выполним осреднение по множеству. Так как высоты последовательных импульсов не связаны корреляционной зависимостью, то среднее по множеству от произведений $(a_k - \bar{a})(a_l - \bar{a})$ равно нулю, за исключением случая $k = l$. При $k = l$ среднее по множеству от $(a_k - \bar{a})(a_l - \bar{a})$ равно $\overline{a^2} - (\bar{a})^2$. Таким образом, предел, входящий в правую часть равенства (9.36), равен

$$\lim_{N\to\infty} \frac{1}{2N} \sum_{-N}^{N} \sum_{-N}^{N} \overline{(a_k - \bar{a})(a_l - \bar{a})}\, e^{-i\omega(k-l)T} = \overline{a^2} - (\bar{a})^2.$$

Средние по множеству от $a_k - \bar{a}$ и $a_l - \bar{a}$ равны, как легко видеть, нулю. Поэтому окончательно будем иметь

$$\Phi(\omega) = \frac{4\pi}{T} |\alpha(\omega)|^2 \left\{ [\overline{a^2} - (\bar{a})^2] + (\bar{a})^2 \lim_{N\to\infty} \frac{1}{2N} \left| \sum_{-N}^{N} e^{-i\omega kT} \right|^2 \right\}. \tag{9.37}$$

Если $\omega = 2\pi n/T$, где n — целое число, то сумма в правой части (9.37) равна $2N+1$. Отсюда следует, что в пределе правая часть равенства (9.37) обратится в бесконечность. При других значениях ω эта сумма будет конечной, и при $N \to \infty$ ее предельное значение будет равно нулю. Поэтому становится ясным, что в пределе правая часть (9.37) изображает последовательность пиков, или δ-функций, повторяющихся с частотой $2\pi n/T$. Для определения коэффициентов при δ-функциях нам нужно подсчитать площадь, ограниченную этой кривой на типовом интервале $-\pi < \omega T < \pi$. Интегрируя сумму, находим по общему правилу величину этой площади:

$$\int_{-\frac{\pi}{T}}^{\frac{\pi}{T}} d\omega \frac{1}{2N} \left| \sum_{-N}^{N} e^{-i\omega kT} \right|^2 = \frac{1}{2N} \int_{-\frac{\pi}{T}}^{\frac{\pi}{T}} \frac{1-\cos 2N\omega T}{1-\cos \omega T}\, d\omega = \frac{2\pi}{T}.$$

Но, в силу соотношений (9.25), площадь, охватываемая кривой, соответствующей δ-функции, равна единице, поэтому искомый коэффициент равен $2\pi/T$ и, следовательно, спектральная плотность рассматриваемой стационарной

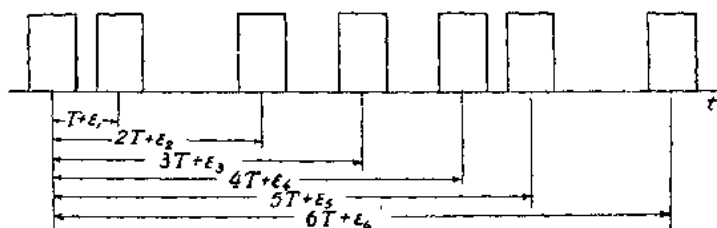
случайной функции определяется соотношением

$$\Phi(\omega) = 2\omega_0 |\alpha(\omega)|^2 \left\{ [\overline{a^2} - (\overline{a})^2] + (\overline{a})^2 \omega_0 \sum_{n=0}^{\infty} \delta(\omega - n\omega_0) \right\}, \tag{9.38}$$

где ω_0 — частота, соответствующая основному периоду T, т. е.

$$\omega_0 = \frac{2\pi}{T}. \tag{9.39}$$

Отсюда следует, что спектральная плотность содержит непрерывную часть, имеющую ту же форму, что и спектральная плотность единичного импульса. Интенсивность



Фиг. 63

этой непрерывной части определяется квадратом среднего отклонения σ высоты импульсов. Кроме того, имеется дискретный спектр с частотами $n\omega_0$ (n — целое число), интенсивность которого также определяется спектром единичного импульса.

Перейдем к рассмотрению последовательности импульсов, имеющих одинаковую форму и одинаковую высоту, но в условиях изменения периода их повторения около его среднего значения. В этих условиях промежуток времени между наступлением одинаковых фаз двух последовательных импульсов равен $T + \varepsilon$, где ε определяется заданным законом $P(\varepsilon)$ распределения вероятностей. Среднее значение ε, по предположению, равно нулю. Что касается последовательных значений ε, то мы вновь предположим,

что между ними не имеется никакой корреляции. Такая случайная функция с прямоугольными импульсами изображена на фиг. 63. Поэтому случайная функция такого типа описывается уравнением

$$y(t) = \sum_k \eta(t - kT - \varepsilon_k), \tag{9.40}$$

где $\eta(t)$ представляет собой одинарный импульс. В силу равенства (9.18), при $\theta = 2NT$

$$A(\omega) = \alpha(\omega) \sum_{-N}^{N} e^{-i\omega(kT + \varepsilon_k)},$$

где $\alpha(\omega)$ представляет собой спектр Фурье для одинарного импульса, описываемый соотношениями (9.34) и (9.35). Согласно равенствам (9.19) и (9.20), спектральная плотность функции $y(t)$ имеет вид

$$\Phi(\omega) = \frac{4\pi}{T} |\alpha(\omega)|^2 \lim_{N\to\infty} \frac{1}{2N} \left[\sum_{-N}^{N} \sum_{-N}^{N} e^{-i\omega\varepsilon_k} e^{+i\omega\varepsilon_l} e^{-i\omega(k-l)T} \right]. \tag{9.41}$$

Введем теперь функцию $\chi(\omega)$ посредством соотношения

$$\chi(\omega) = \int_{-\infty}^{\infty} P(\varepsilon) e^{-i\omega\varepsilon} d\varepsilon. \tag{9.42}$$

Функцию $\chi(\omega)$ иногда называют характеристической функцией величины ε; она представляет собой изображение Фурье функции $P(\varepsilon)$[1]. Мы можем написать теперь

$$e^{-i\omega\varepsilon_k} e^{+i\omega\varepsilon_l} =$$
$$= \{[e^{-i\omega\varepsilon_k} - \chi(\omega)] + \chi(\omega)\} \{[e^{+i\omega\varepsilon_l} - \chi^*(\omega)] + \chi^*(\omega)\},$$

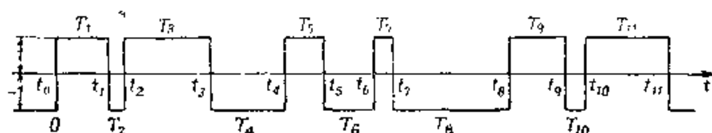
где величина $\chi^*(\varepsilon)$ означает функцию, комплексно сопряженную с $\chi(\omega)$, и равна $\chi(-\omega)$. Подставим теперь правую часть последнего равенства в выражение, стоящее под знаком двойной суммы в правой части (9.41), и прежде всего осуществим осреднение по множеству. При этом

[1]) Автор книги иногда, но не всегда, пользуется множителем $1/2\pi$ перед интегралом в формуле для изображения Фурье; см. формулы (9.18) и (9.42). — *Прим. ред.*

процесс предельного перехода сильно облегчается тем обстоятельством, что значения величин ε_k не связаны никакой корреляцией. Окончательное выражение функции $\Phi(\omega)$ имеет вид

$$\Phi(\omega) = 2\omega_0 |\alpha(\omega)|^2 \{[1 - |\chi(\omega)|^2] + |\chi(\omega)|^2 \omega_0 \sum_{n=0}^{\infty} \delta(\omega - n\omega_0)\}, \tag{9.43}$$

где ω_0 — частота, определенная соотношением (9.39). В данном случае характер непрерывного и дискретного



Фиг. 64

спектров уже не определяется только спектром одинарного импульса, а зависит также и от характеристической функции величины ε.

В качестве третьего примера прямого метода подсчета спектральной функции рассмотрим стационарную случайную функцию $y(t)$, изображенную на фиг. 64. Эта функция может принимать только два значения: $+1$ и -1, чередующиеся через некоторый интервал T. Однако T не является постоянной величиной, а определяется ее вероятным распределением $P(T)$, причем $T > 0$. Примем также, что между последовательными интервалами T нет никакой корреляции. Последовательные интервалы обозначим через T_k, где $k = 1, 2, \ldots$, и примем за период Θ, по которому выполняется интегрирование в правой части (9.18), промежуток времени от $t = 0$ до $t = N\bar{T}$, где $\bar{T}$ — среднее значение интервала времени, определяемое интегралом

$$\bar{T} = \int_0^{\infty} TP(T)\, dT. \tag{9.44}$$

Тогда

$$A(\omega) = \frac{1}{2\pi} \int_0^{N\bar{T}} y(t) e^{-i\omega t}\, dt = \frac{1}{2\pi} \frac{1}{i\omega} \sum_{k=1}^{N} (-1)^k (e^{-i\omega t_k} - e^{-i\omega t_{k-1}}).$$

Здесь t_k — момент времени, соответствующий концу k-го интервала. Последнее выражение можно преобразовать к виду

$$A(\omega) = \left[\frac{1}{\pi}\frac{1}{i\omega}\sum_{k=1}^{N}(-1)^k e^{-i\omega t_k}\right] - \frac{1}{2\pi}\frac{1}{i\omega}(-1)^N e^{-i\omega t_N} + 1.$$

С помощью равенства (9.20) и последнего соотношения мы получаем спектральную плотность $\Phi(\omega)$ в виде

$$\Phi(\omega) = \frac{4}{\pi \bar{T}\omega^2}\left[\lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N}\sum_{k'=1}^{N}(-1)^{k+k'} e^{-i\omega(t_k - t_{k'})}\right]. \quad (9.45)$$

Рассмотрим теперь некоторое $k > k'$, например $k = k' + m$; тогда

$$e^{-i\omega(t_k - t_{k'})} = e^{-i\omega T_{k'+1}} e^{-i\omega T_{k'+2}} \ldots e^{-i\omega T_{k'+m}}. \quad (9.46)$$

Так как между последовательными интервалами не имеется никакой корреляции, то вероятность появления произведения, выписанного в правой части (9.46), равна произведению вероятностей появления величин, равных сомножителям этого произведения. Введя характеристическую функцию $\chi(\omega)$ распределения $P(T)$

$$\chi(\omega) = \varphi(\omega) - i\psi(\omega) = \int_0^\infty P(T)\, e^{-i\omega T}\, dT, \quad (9.47)$$

мы увидим, что $\chi(\omega)$ представляет собой среднее значение $e^{-i\omega T}$. Поэтому среднее по множеству значение произведения в правой части (9.46) равно просто $[\chi(\omega)]^m$. Под знаком двойной суммы в правой части (9.45) имеется приблизительно N таких произведений; все они входят со знаком, определяемым значением $(-1)^m$. Следовательно, в пределе эти произведения сводятся к величине $(-1)^m[\chi(\omega)]^m$. Под знаком предела m изменяется от 1 до ∞ и, следовательно, сумма всех таких членов равна

$$\sum_{m=1}^{\infty}(-1)^m[\chi(\omega)]^m = -\frac{\chi(\omega)}{1+\chi(\omega)}.$$

Члены, соответствующие $k' > k$, равны величинам, комплексно сопряженным с членами, соответствующими $k > k'$.

Помимо них, в двойной сумме в правой части равенства (9.45) содержатся еще члены, соответствующие значениям $k=k'$; эти члены дают в сумме, после деления на N, единицу. Таким образом, равенство (9.45) окончательно принимает вид

$$\Phi(\omega)=\frac{4}{\pi \bar{T}\omega^2}\left\{1-2\,\mathrm{Re}\left[\frac{\chi(\omega)}{1+\chi(\omega)}\right]\right\}.$$

Обозначим через $\varphi(\omega)$ и $\psi(\omega)$ вещественную и мнимую части функции $\chi(\omega)$ соответственно, как это принято в соотношении (9.47). Тогда

$$\Phi(\omega)=\frac{4}{\pi \bar{T}\omega^2}\,\frac{1-\varphi^2(\omega)-\psi^2(\omega)}{[1+\varphi(\omega)]^2+\psi^2(\omega)}. \tag{9.48}$$

Если $P(T)$ определяется распределением Пуассона

$$P(T)=\frac{1}{\bar{T}}e^{-T/\bar{T}}, \tag{9.49}$$

где $\bar{T}$ — средний интервал времени (9.44), то спектральная плотность такой случайной функции единичной амплитуды равна

$$\Phi(\omega)=\frac{\bar{T}}{\pi}\,\frac{1}{1+\left(\frac{\omega\bar{T}}{2}\right)^2}. \tag{9.50}$$

Благодаря полному отсутствию какой-бы то ни было установившейся периодичности в рассматриваемой случайной функции ее спектральная плотность является непрерывной и имеет гладкий характер; в ней не имеется никаких пиков, наблюдавшихся в предыдущих примерах.

9.6. Вероятность больших отклонений от среднего значения. Если случайная функция характеризует напряжение в материале конструкции, то недостаточно знать только среднее значение напряжения. В целях обеспечения прочности конструкции мы должны разыскать также и вероятность возникновения напряжений, превышающих допустимое рабочее напряжение в материале конструкции, т. е. вероятность $P[\,|y|>k]$ того, что случайная функция y примет значение, превосходящее величину k. Эта вероятность находится очень легко, если известна первая

функция распределения вероятностей $W_1(y)$:

$$P[|y| > k] = \int_{-\infty}^{-k} W_1(y)\,dy + \int_{k}^{\infty} W_1(y)\,dy. \qquad (9.51)$$

Но во многих технических задачах эта функция распределения вероятностей не задана, известны же среднее значение y и среднее отклонение σ. Однако даже и при этих ограниченных данных все же возможно дать общую, хотя и грубую оценку вероятности возникновения больших отклонений от среднего значения. Если, например, $g(y)$ — неотрицательная функция от y, то, так как функция $W_1(y)$ неотрицательна по определению,

$$\overline{g(y)} = \int_{-\infty}^{\infty} g(y)\,W_1(y)\,dy > K \int_{g(y)>K} W_1(y)\,dy. \qquad (9.52)$$

При вычислении последнего интеграла интегрирование распространяется на все значения y, удовлетворяющие условию $g(y) > K$. Но этот интеграл как раз равен $P[g(y) > K]$. Таким образом,

$$P[g(y) \geqslant K] \leqslant \frac{\overline{g(y)}}{K}. \qquad (9.53)$$

Последнее неравенство называется неравенством Чебышева. Положим теперь

$$g(y) = (y - \overline{y})^2.$$

Тогда, в силу (9.7),

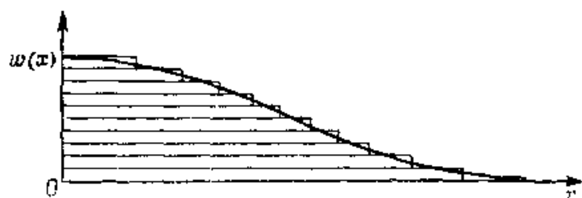
$$\overline{g(y)} = \sigma^2 = \overline{y^2} - (\overline{y})^2,$$

где σ — среднее отклонение от среднего значения. Введя постоянную K посредством соотношения $K = k^2\sigma^2$, мы получим с помощью неравенства (9.53) неравенство Бьенэмэ — Чебышева:

$$P[|y - \overline{y}| \geqslant k\sigma] \leqslant \frac{1}{k^2}. \qquad (9.54)$$

Как известно, это неравенство является слишком грубым для большинства прикладных задач; определяемая им верхняя граница в общем случае оказывается чрез-

мерно завышенной. Более точную оценку можно дать в тех случаях, когда функция $W_1(y)$ имеет единственный максимум. Соответствующая оценка принадлежит Гауссу. Для доказательства неравенства Гаусса рассмотрим функцию $w(x)$ (фиг. 65), монотонно убывающую при $x > 0$.



Фиг. 65

Очевидно, что $w(x)$ можно представлять себе как функцию, образованную путем наложения функций, постоянных (не равных нулю) от $x=0$ до $x=x_0$ и равных нулю при $x > x_0$. Пусть $v(x)=1$ при $0 < x \leqslant x_0$ и $v(x)=0$ при $x > x_0$. Тогда при любом $K > x_0$

$$K^2 \int_K^\infty v(x)\,dx = 0.$$

Но, коль скоро $0 < K < x_0$,

$$K^2 \int_K^\infty v(x)\,dx = K^2(x_0 - K).$$

При значениях K, заключенных в указанных пределах, максимум этой величины равен $\frac{4}{27}x_0^3$. Таким образом, в общем случае справедливо неравенство

$$K^2 \int_K^\infty v(x)\,dx \leqslant \frac{4}{9}\int_0^\infty x^2 v(x)\,dx.$$

Переходя от $v(x)$ к $w(x)$ по методу наложения, как указано выше, получим

$$K^2 \int_K^\infty w(x)\,dx \leqslant \frac{4}{9}\int_0^\infty x^2 w(x)\,dx.$$

Рассмотрим теперь произвольное распределение с единственным максимумом и с абсциссой $x = y - y_0$, где y_0 — максимум. Тогда

$$K^2 \int_K^\infty W_1(x)\,dx \leqslant \frac{4}{9} \int_0^\infty x^2 W_1(x)\,dx$$

$$K^2 \int_{-\infty}^{K} W_1(x)\,dx \leqslant \frac{4}{9} \int_{-\infty}^{0} x^2 W_1(x)\,dx.$$

Складывая эти выражения, получаем

$$K^2 P[|y - y_0| \geqslant K] \leqslant \frac{4}{9} \nu^2,$$

где ν — среднее отклонение от максимума, определяемое равенством

$$\nu^2 = \overline{(y - y_0)^2} = \int_{-\infty}^{\infty} (y - y_0)^2 W_1(y)\,dy. \qquad (9.55)$$

Положив $K = k\nu$, придем к неравенству Гаусса

$$P[|y - y_0| \geqslant k\nu] \leqslant \frac{4}{9k^2}. \qquad (9.56)$$

Если распределение симметрично, то $y_0 = \bar{y}$, $\nu = \sigma$ и неравенства (9.56) сводятся к неравенствам

$$P[|y - \bar{y}| \geqslant k\sigma] \leqslant \frac{4}{9k^2}. \qquad (9.57)$$

Неравенства (9.57) доставляют лучшую оценку вероятности по сравнению с неравенствами (9.54).

В реальных условиях часто возможно предполагать гауссово распределение вероятностей, по крайней мере приближенно. Тогда, пользуясь асимптотическим разложением функции ошибки, легко показать, что

$$P[|y - \bar{y}| \geqslant k\sigma] \approx \frac{e^{-\frac{1}{2}k^2}}{k\sqrt{2\pi}} \quad \text{при} \quad k \gg 1. \qquad (9.58)$$

Это очень малая вероятность. Например, при $k = 3$ вероятность равна всего 0,002. Пользуясь же оценкой (9.54),

мы сможем утверждать лишь то, что вероятность меньше 0,1111, тогда как оценка (9.57) означает, что вероятность меньше 0,0493. Различие в качестве этих оценок объясняется, конечно, различием в количестве информации, на которой основана оценка.

9.7. Частота появления значений, превышающих некоторую заданную величину. Если случайная функция характеризует напряжение в материале конструкции и если расчет этой конструкции на прочность должен основываться на повторениях определенной величины напряжения, т. е. на усталости материала, то необходимо знать вероятное число превышения случайной функцией значения $y=\xi$ в единицу времени. Очевидно, что это число равно половине числа прохождений случайной функции через значение ξ в единицу времени. Обозначим его через $N_0(\xi)$. Это число впервые подсчитал С. О. Райс; мы будем следовать его методу[1]).

Пусть $W(y, y')\,dy\,dy'$ есть вероятность того, что случайная функция принимает значение, находящееся между y и $y+dy$ в то же время, когда ее производная y' по времени принимает значение, находящееся между y' и $y'+dy'$. Эту вероятность можно также рассматривать как отношение промежутка времени, в течение которого случайная функция y и ее производная y' одновременно заключены в указанных интервалах, к некоторой единице времени. Но функция y пробегает интервал dy за время $dy/|y'|$. Отсюда следует, что ожидаемое или вероятное число прохождений через ξ и y' в единицу времени равно отношению величин $W(\xi, y')\,dy\,dy'$ и $dy/|y'|$, т. е. $|y'|\,W(\xi, y')\,dy'$. Искомое число $N_0(\xi)$ найдется в результате интегрирования по всем значениям y', т. е.

$$N_0(\xi) = \int_{-\infty}^{\infty} |y'|\,W(\xi, y')\,dy'. \tag{9.59}$$

Но, как следует из (9.15), для любой дифференцируемой случайной функции между y и y' нет никакой корреляции.

---

[1]) Rice S. O., Bell System Tech. Journ., 23, 282 (1944); 25, 46 (1945).

Тогда, в силу общих законов теории вероятностей, $W(y, y')$ просто равно произведению первых функций распределения вероятностей $W_1(y)$ и $W(y')$. Следовательно, равенство (9.59) можно написать в виде

$$N_0(\xi) = W_1(\xi) \int_{-\infty}^{\infty} |y'| W(y')\, dy'. \qquad (9.60)$$

Если распределение $W(y')$ симметрично, то равенство (9.60) можно еще свести к виду

$$N_0(\xi) = 2W_1(\xi) \int_0^{\infty} y' W(y')\, dy' \text{ при } W(y') \text{ симметричном.} \qquad (9.61)$$

Если $W(y')$ определяет гауссово распределение со средним отклонением σ', то, в силу (9.9),

$$N_0(\xi) = \frac{2W_1(\xi)}{\sigma'\sqrt{2\pi}} \int_0^{\infty} y' e^{-\frac{y'^2}{2\sigma'^2}}\, dy' = \frac{2\sigma' W(\xi)}{\sqrt{2\pi}}. \qquad (9.62)$$

Среднее отклонение σ' можно подсчитать, исходя из спектральной плотности $\Phi(\omega)$ и пользуясь равенствами (9.15) и (9.22):

$$\sigma'^2 = \int_0^{\infty} \omega^2 \Phi(\omega)\, d\omega. \qquad (9.63)$$

Если $W_1(y)$ также представляет собой гауссово распределение со средним значением $\bar{y}$ и средним отклонением σ, то, принимая во внимание (9.7) и (9.21), получаем

$$N_0(\xi) = \frac{1}{\pi} \frac{\sigma'}{\sigma} e^{-\frac{1}{2}\frac{(\xi-\bar{y})^2}{\sigma^2}} = \frac{1}{\pi} e^{-\frac{1}{2}\frac{(\xi-\bar{y})^2}{\sigma^2}} \left[\frac{\int_0^{\infty} \omega^2 \Phi(\omega)\, d\omega}{\int_0^{\infty} \Phi(\omega)\, d\omega - (\bar{y})^2} \right]^{\frac{1}{2}}. \qquad (9.64)$$

Последнее равенство и представляет собой формулу **Райса.**

9.8. Реакция линейной системы на стационарное случайное входное воздействие. Ответим, наконец, на вопрос, поставленный во введении к этой главе: как подсчитать выход линейной системы с постоянными параметрами, на вход которой действует стационарная случайная функция?

В свете элементов теории случайных функций, изложенных в предыдущих разделах, становится очевидным, что ключ к ответу на этот вопрос лежит в подсчете спектральной плотности выхода по спектральной плотности входа. Если же спектральная плотность выхода найдена, то нетрудно подсчитать с помощью (9.22) функцию корреляции и с помощью (9.21) среднее квадратичное значение. После этого, применяя методы, описанные в п.п. 9.6 и 9.7, можно оценить вероятность возникновения больших отклонений от среднего значения и частоту появления величин, превышающих некоторую заданную величину.

Для многих технических задач в общем случае достаточно знать именно эти характеристики выходной переменной системы.

Пусть на вход системы подана случайная функция $x(t)$, обладающая спектральной плотностью $\Phi(\omega)$ и функцией корреляции $R_i(\tau)$. Тогда, в силу (9.21) и (9.22), мы будем иметь

$$\overline{x^2} = \int_0^\infty \Phi(\omega)\,d\omega = R_i(0) \tag{9.65}$$

и

$$R_i(\tau) = \int_0^\infty \Phi(\omega)\cos\omega\tau\,d\tau = \frac{1}{2}\int_{-\infty}^{\infty} \Phi(\omega)\,e^{i\omega\tau}\,d\omega, \tag{9.66}$$

где, как это следует из соотношения (9.23), используется равенство $\Phi(-\omega) = \Phi(\omega)$. Подобным же образом, обозначив через $y(t)$ выход системы, а через $g(\omega)$ — его спектральную плотность и через $R_0(\tau)$ — его функцию корреляции, получим, что

$$\overline{y^2} = \int_0^\infty g(\omega)\,d\omega = R_0(0) \tag{9.67}$$

и

$$R_0(\tau) = \frac{1}{2} \int_{-\infty}^{\infty} g(\omega) e^{i\omega\tau} d\omega. \quad (9.68)$$

Пусть, как и выше, $h(t)$ означает реакцию линейной системы на единичный импульс, приложенный к системе в момент времени $t=0$. Если процесс начинается с момента времени $t=-\infty$, то выход системы можно записать в виде

$$y(t) = \int_{-\infty}^{t} x(\tau) h(t-\tau) d\tau,$$

где $x(\tau)d\tau$ представляет собой импульс, приложенный в момент $t=\tau$. Введем новую переменную интегрирования $u=t-\tau$. Тогда

$$y(t) = \int_0^{\infty} x(t-u) h(u) du. \quad (9.69)$$

Таким образом, корреляционная функция $R_0(\tau)$ определяется выражением

$$R_0(\tau) = \overline{y(t) y(t+\tau)} = \int_0^{\infty} \int_0^{\infty} \overline{x(t-u) x(t+\tau-u')} h(u) h(u') du du'.$$

Но

$$\overline{x(t-u) x(t+\tau-u')} = \overline{x(t) x(t+\tau+u-u')} = R_i(\tau+u-u').$$

Следовательно, применяя соотношения (9.66) и (9.68), получим

$$\int_{-\infty}^{\infty} g(\omega) e^{i\omega\tau} d\omega = \int_{-\infty}^{\infty} \int_0^{\infty} \int_0^{\infty} \Phi(\omega) e^{i\omega(\tau+u-u')} h(u) h(u') du du' d\omega. \quad (9.70)$$

Введем теперь передаточную функцию $F(s)$ линейной системы, представляющую собой изображение по Лапласу функции $h(t)$. Обращаясь к соотношению (3.50), напишем

$$F(i\omega)=\int_0^\infty e^{-i\omega u}h(u)\,du.$$

Принимая во внимание это равенство, преобразуем соотношение (9.70) к виду

$$\int_{-\infty}^{\infty} g(\omega)e^{i\omega\tau}\,d\omega=\int_{-\infty}^{\infty}\Phi(\omega)F(i\omega)F(-i\omega)e^{i\omega\tau}\,d\omega.$$

В результате мы получим соотношение, связывающее спектральные плотности $g(\omega)$ и $\Phi(\omega)$:

$$g(\omega)=\Phi(\omega)F(i\omega)F(-i\omega)=|F(i\omega)|^2\,\Phi(\omega), \tag{9.71}$$

где принято во внимание, что функции $F(i\omega)$ и $F(-i\omega)$ представляют собой комплексные сопряженные величины.

Соотношение (9.71) определяет спектральную плотность выхода по спектральной плотности входа и частотной характеристике данной линейной системы. С помощью этого соотношения спектральную плотность $g(\omega)$ легко подсчитать даже в тех случаях, когда частотная характеристика задана в форме графика или таблицы числовых значений. Это еще раз указывает на важность понятия передаточной функции и частотной характеристики. Интересно отметить, что, так как функция $F(i\omega)$ обычно обращается в нуль при $\omega\to\infty$, спектральная плотность $g(\omega)$ выхода стремится к нулю при $\omega\to\infty$ быстрее спектральной плотности $\Phi(\omega)$ входа. Это обстоятельство приводит к «сглаживанию» выходной функции.

9.9. Система второго порядка. Обратимся к простому примеру, основанному на рассмотрении системы второго порядка. Движение этой системы описывается уравнением

$$m\frac{d^2y}{dt^2}+c\frac{dy}{dt}+ky=x(t). \tag{9.72}$$

Следовательно, передаточная функция системы имеет вид

$$F(s) = \frac{1}{ms^2 + cs + k} = \frac{1}{k}\,\frac{1}{(s/\omega_0)^2 + 2\zeta(s/\omega_0) + 1},$$

где ω_0 — частота собственных колебаний системы без демпфирования, а ζ — отношение истинного коэффициента демпфирования к критическому значению этого коэффициента, определенное соотношением (3.38). Поэтому

$$F(i\omega)F(-i\omega) = \frac{1}{k^2\{[(\omega/\omega_0)^2 - 1]^2 + 4\zeta^2(\omega/\omega_0)^2\}}.$$

Следовательно, спектральная плотность выхода определяется соотношением

$$g(\omega) = \frac{\Phi(\omega)}{k^2\{[(\omega/\omega_0)^2 - 1]^2 + 4\zeta^2(\omega/\omega_0)^2\}}. \tag{9.73}$$

С помощью (9.21) мы найдем среднее квадратичное значение выхода рассматриваемой линейной системы:

$$\overline{y^2} = \frac{1}{k^2}\int_0^\infty \frac{\Phi(\omega)\,d\omega}{[(\omega/\omega_0)^2 - 1]^2 + 4\zeta^2(\omega/\omega_0)^2}. \tag{9.74}$$

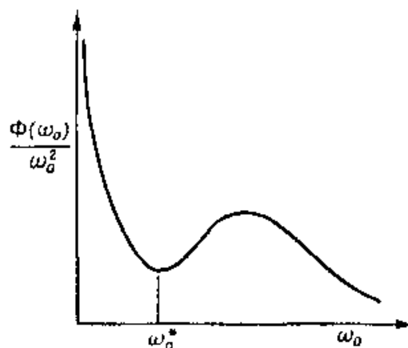
Заметим теперь, что если ζ достаточно мало, то при $\omega = \omega_0$ знаменатель подинтегральной функции в правой части равенства (9.74) приближается к нулю. Поэтому, если $\Phi(\omega)$ — функция, изменяющаяся медленно, то

$$\overline{y^2} \approx \frac{\omega_0\Phi(\omega_0)}{k^2}\int_0^\infty \frac{dx}{(x^2 - 1)^2 + 4\zeta^2x^2} =$$

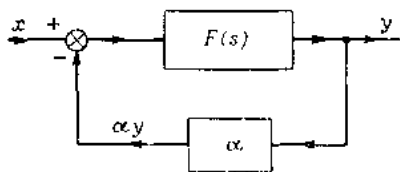
$$= \frac{1}{k^2}\,\omega_0\Phi(\omega_0)\frac{\pi}{4\zeta} = \frac{\pi}{2mc}\,\frac{\Phi(\omega_0)}{\omega_0^2}. \tag{9.75}$$

Это соотношение показывает, что если коэффициент c обращается в нуль, то среднее квадратичное значение выхода становится бесконечно большим. При $c = 0$ передаточная функция имеет чисто мнимый полюс в точке $i\omega_0$. В общем случае это явление неограниченного возрастания выхода имеет место каждый раз, когда передаточная функция линейной системы имеет чисто мнимый полюс. Поэтому, чтобы система действовала удовлетворительным

образом при подаче на ее вход случайной функции, на передаточную функцию системы необходимо наложить следующее требование: все полюсы передаточной функции $F(s)$ должны иметь отрицательные вещественные части. Таким образом, это основное требование, налагаемое на систему при входных функциях случайного вида, оказывается тождественным требованию, обеспечивающему устойчивость системы при входных функциях обычного вида[1]).



Фиг. 66



Фиг. 67

С помощью дополнительных изменений передаточной функции системы можно в общем случае обеспечить дальнейшее улучшение качества процесса на выходе системы. Например, вполне очевидно, что функция $\Phi(\omega_0)/\omega_0^2$, входящая в правую часть (9.75), достигает минимума при некоторой частоте ω_0^* (фиг. 66). В этих условиях имеется возможность добиться уменьшения случайных флуктуаций на выходе, обеспечив эффективную отработку системы при $\omega = \omega_0^*$. Для этого достаточно ввести в систему простую пропорциональную обратную связь, показанную на фиг. 67. Тогда движение системы с обратной связью будет описываться уже не уравнением (9.72), а уравнением

$$m\frac{d^2y}{dt^2} + c\frac{dy}{dt} + ky = x - \alpha y$$

[1]) Физически это свойство объясняется тем, что при приближении системы к идеальному резонансному контуру с собственной частотой ω_0, компонента частотного спектра, имеющая частоту ω_0 и содержащаяся во входной функции, может «раскачать» колебания в системе до сколь угодно большой амплитуды. — *Прим. ред.*

или

$$m\frac{d^2y}{dt^2}+c\frac{dy}{dt}+(k+\alpha)y=x. \qquad (9.76)$$

Чтобы сделать собственную частоту системы равной ω_0^*, достаточно потребовать выполнения соотношения

$$\omega_0^{*2}=\frac{k+\alpha}{m}. \qquad (9.77)$$

Поэтому путем надлежащего выбора коэффициента α обратной связи оказывается возможным уменьшить выходную переменную системы.

9.10. Подъемная сила двумерного аэродинамического профиля, обтекаемого турбулентным потоком несжимаемой жидкости. В качестве второго примера рассмотрим тонкий плоский аэродинамический профиль с хордой c, перемещающийся с постоянной скоростью U в турбулентном потоке воздуха. Направим ось x вдоль хорды, ось z — по размаху крыла и ось y — перпендикулярно к направлениям хорды и размаха[1]). Предположим, что составляющие u, v и w турбулентных флуктуаций скорости малы по сравнению с U. Благодаря влиянию этих турбулентных флуктуаций угол атаки будет меняться во времени и, следовательно, возникнут флуктуации подъемной силы профиля. Возмущение α угла атаки определяется отношением

$$\alpha=\frac{v}{U},$$

справедливым при малых значениях флуктуаций[2]). Здесь функция $\alpha(t)$ играет роль вынуждающей функции. «Реакция» же системы определяется подъемной силой профиля с учетом флуктуаций или, лучше сказать, определяется величиной коэффициента $C_l(t)$ подъемной силы с учетом

---

[1]) См. также правый рисунок на фиг. 27. — *Прим. перев.*

[2]) Так как в данной системе осей $\alpha=\operatorname{arctg}\frac{v}{U+u}\approx\frac{v}{U}$. — *Прим. перев.*

флуктуаций[1]). Такая задача была рассмотрена Г. У. Липманом[2]).

Для подсчета среднего квадратичного значения $\overline{C_l^2}(t)$ коэффициента подъемной силы необходимо, прежде всего, найти передаточную функцию данного аэродинамического профиля. Но это уже было выполнено в п. 3.7. В самом деле, частотная характеристика $F(i\omega)$ такого профиля определена с помощью соотношений (3.54)–(3.58), в которых входом служит флуктуация v, а выходом — коэффициент C_l подъемной силы.

Следует иметь в виду, что флуктуации, обусловленные возникновением турбулентности, являются трехмерными по самому существу явления турбулентности, т. е. каждая из переменных u, v и w зависит от переменных x, y, z и t. В первом же приближении представляется достаточным учитывать только составляющую v и ее зависимость от x и t. Таким образом, мы считаем, что турбулентность потока порождает флуктуацию v скорости или, что то же, флуктуацию угла атаки, определяемую отношением

$$\alpha(x, t) = \frac{v(x, t)}{U}.$$

Предположим теперь, что характер турбулентности не претерпевает заметных изменений на протяжении промежутка времени порядка c/U. Тогда составляющая угла атаки, обусловленная турбулентностью, будет также зависеть только от разности $t - x/U$, и мы можем применить выводы Сирса, приведенные в п. 3.7. Такое предположение часто вводится при исследовании турбулентности, и оно существенно основано на выполнении следующего условия: быстрота изменения скорости отдельных жидких частиц при перемещении вдоль их траекторий мала по сравнению с быстротой изменения скорости

---

[1]) Подъемная сила $\Pi = \frac{S\rho}{2} V^2 C_l(\alpha)$, где $\frac{S\rho}{2}$ здесь постоянно, а $V = \sqrt{(U+u)^2 + v^2 + w^2} \approx U$ при малых u, v, w. Поэтому здесь можно перейти от Π к C_l. — *Прим. перев.*

[2]) Liepmann H. W., Journ. Aeronaut. Sci., 19, 793—801 (1952).

потока в фиксированных точках жидкой среды[1]). При этом предположении

$$\overline{C_l^2} = 4\pi^2 \int_0^\infty \Phi(\omega) |\varphi(k)|^2 d\omega, \tag{9.78}$$

где $\Phi(\omega)$ — спектральная плотность случайной функции v/U.

В силу (9.31) и (9.32),

$$\Phi(\omega) = \frac{\overline{v^2}}{U^2} \frac{L}{\pi U} \frac{1 + 3\frac{L^2\omega^2}{U^2}}{\left[1 + \frac{L^2\omega^2}{U^2}\right]^2}. \tag{9.79}$$

Далее, Липман обнаружил, что функцию $|\varphi(k)|^2$ можно аппроксимировать выражением

$$|\varphi(k)|^2 \approx \frac{1}{1 + 2\pi k}. \tag{9.80}$$

Таким образом,

$$\overline{C_l^2} = 4\pi^2 \frac{\overline{v^2}}{U^2} \int_0^\infty \frac{1 + 3u^2}{(1 + u^2)^2} \frac{1}{1 + \eta u} du =$$

$$= 4\pi^2 \frac{\overline{v^2}}{U^2} \left[\frac{4\eta - \pi}{2\pi(\eta^2 + 1)} + \frac{\eta^2 + 3}{2\pi(\eta^2 + 1)^2} (\eta \ln \eta^2 + \pi) \right], \tag{9.81}$$

где

$$\eta = \frac{\pi c}{L}. \tag{9.82}$$

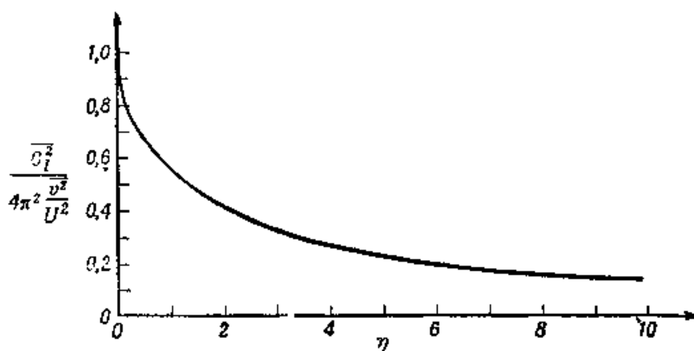
Найденная зависимость среднего квадратичного значения коэффициента подъемной силы от параметра η графически изображена на фиг. 68.

Очевидно, что при $c/L \to 0$ мы приходим к случаю движения аэродинамического профиля малой хорды в сильно турбулентном потоке и

$$\overline{C_l^2} \to 4\pi^2 \frac{\overline{v^2}}{U^2} = 4\pi^2 \overline{\alpha^2}.$$

[1]) Этот способ оценки скоростей существенно связан с методами описания движения жидкой среды по Лагранжу, с одной стороны, и по Эйлеру — с другой, см. [35]. — *Прим. перев.*

Таким образом, обтекание профиля носит квазистационарный характер, и наклон кривой C_l по α равен 2π. С другой стороны, при больших значениях c/L мы приходим к случаю профиля длинной хорды, движущегося в потоке слабой турбулентности. Из соотношения (9.80) вытекает, что в этом случае $\overline{C_l^2} \to 0$. Это означает, что «порывы»,



Фиг. 68

возмущающие течение, полностью компенсируют друг друга и чистая подъемная сила равна нулю[1]) при образовании ударных перегрузок крыла.

9.11. Пульсирующий вход. Очень важное значение имеет явление «пульсирования» в части потока, являющейся аэродинамическим следом. Дело в том, что частицы воздуха, находящиеся на границе всякого аэродинамического следа, совершают усиленные флуктуационные колебания таким образом, что всякая частица, расположенная вблизи этой границы, будет находиться то внутри этого следа, то вне его. Для частей хвостового оперения, движущихся вблизи границы аэродинамического следа крыла, летящего при критическом угле по всему размаху крыла или

---

[1]) Напомним, что речь идет о приращении подъемной силы и, следовательно, коэффициента C_l, обусловленном турбулентностью. При ламинарном течении коэффициент подъемной силы профиля определяется установившейся составляющей угла атаки и может, в частности, оказаться равным нулю при соответствующем значении этой составляющей. —*Прим. перев.*

по части крыла, это пульсирующее воздействие может, следовательно, иметь исключительно важное значение, определяя подъемную силу горизонтального оперения.

Чтобы составить себе грубое представление о данном явлении, будем рассматривать поток в области хвостового оперения, как стационарное однородное течение со скосом потока, включаемым и выключаемым в беспорядочно следующие один за другим моменты времени[1]). Такое течение служит, вероятно, хорошей моделью аэродинамического следа крыла при критическом угле атаки. Если допустить, что вероятность этого переключения подчиняется распределению Пуассона, то имеется возможность использовать для решения задачи выражение (9.50) спектральной плотности, внеся в него некоторые изменения: здесь среднее значение отклонения уже равно не единице, а среднему значению $\sqrt{\overline{v^2}}/U$ угла атаки; средняя продолжительность промежутка времени, на протяжении которого вынуждающая функция включена, равна среднему интервалу $\overline{T}$. Таким образом, выражение для спектральной плотности имеет в этом случае вид

$$\Phi(\omega) = \frac{\overline{v^2}}{U^2} \cdot \frac{\overline{T}}{\pi} \cdot \frac{1}{1 + \left(\frac{\omega \overline{T}}{2}\right)^2}. \tag{9.83}$$

Следовательно, среднее квадратичное значение коэффициента подъемной силы приблизительно равно

$$\overline{C_l^2} = \frac{\overline{v^2}}{U^2} \frac{\overline{T}}{\pi} 4\pi^2 \int_0^\infty \frac{d\omega}{\left[1 + \left(\frac{\omega \overline{T}}{2}\right)^2\right]\left[1 + \pi\left(\frac{\omega c}{U}\right)\right]} =$$

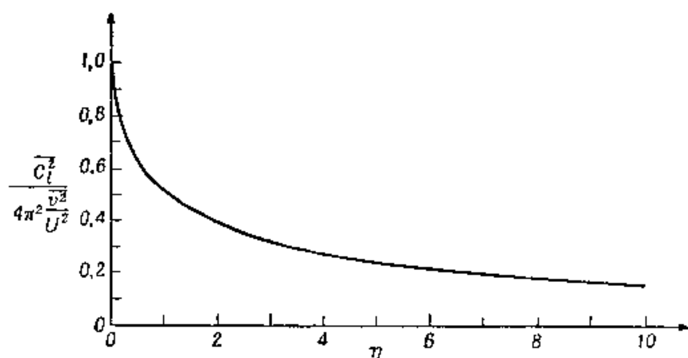
$$= 4\pi^2 \frac{\overline{v^2}}{U^2} \frac{2}{\pi} \frac{\eta \ln \eta + \frac{\pi}{2}}{1 + \eta^2} \tag{9.84}$$

при

$$\eta = \frac{2\pi c}{U\overline{T}}.$$

[1]) О скосе потока см. [24, 49]. —*Прим. перев*

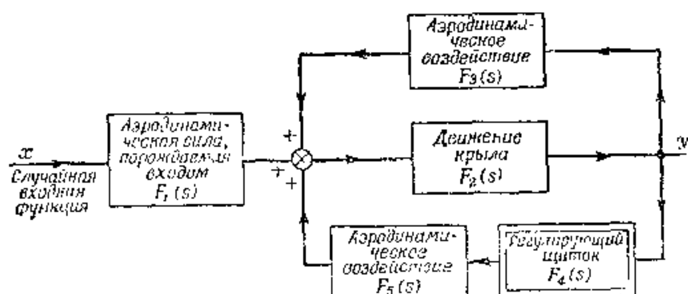
Это соотношение графически изображено на фиг. 69. Предельные значения, отвечающие случаям $\eta = 0$ и $\eta \to \infty$, конечно, совпадают с предельными значениями, рассмотренными в предыдущем пункте.



Фиг. 69

9.12. Методика разработки следящих систем, входом которых служат случайные функции. При рассмотрении реакции системы второго порядка на случайную входную функцию мы указывали на возможность улучшения процессов в системе путем введения в нее обратной связи и превращения ее тем самым в следящую систему. Однако в разобранном примере цепь обратной связи является довольно примитивной в том смысле, что величина корректирующего воздействия, создаваемого обратной связью, и входная вынуждающая функция суть величины одного и того же порядка. В схемах более практического характера эту цепь обратной связи можно делать более гибкой, с тем чтобы уменьшать потребное корректирующее воздействие. Например, движением крыла в турбулентном потоке можно управлять с помощью системы автоматического регулирования, перемещающей управляющий щиток, который поворачивается около оси, фиксированной относительно крыла. В такой конструкции сила, необходимая для передвижения щитка, может быть ничтожной сравнительно с изменением аэродинамических сил, произведенным в результате этого передвижения. Общая схема мыслимой системы автоматического регулирования для

управления положением щитка изображена на фиг. 70. Входной случайной функцией служит турбулентный поток воздуха. Первая аэродинамическая передаточная функция $F_1(s)$ доставляет связь между турбулентным потоком и подъемной силой, обусловленной этим турбулентным потоком; эта функция приближенно определяется выражением (3.56). Под действием изменений подъемной силы



Фиг. 70

и продольного аэродинамического момента крыло начинает перемещаться в направлении, параллельном вертикали, а также поворачиваться в своей плоскости. Связь между аэродинамическими силами и движением крыла определяется передаточной функцией $F_2(s)$ движения. В результате этого движения крыла возникают два явления. Первое из них заключается в появлении аэродинамических сил, обусловленных движением крыла[1]) и характеризуемых второй аэродинамической передаточной функцией $F_3(s)$. На схеме это явление изображается первой цепью обратной связи, которую, однако, проектировщик не может задать по своему усмотрению. Но в распоряжении проектировщика находится вторая цепь обратной связи, характеризующая второе влияние движения крыла; от этого движения производится управление перемещением регулирующего щитка, согласно передаточной функции $F_4(s)$. В свою очередь щиток, перемещаясь относительно крыла, создает новые аэродинамические силы, характеризуемые передаточной функцией $F_5(s)$. Таким

[1]) Эти силы возникают благодаря изменению характера обтекания крыла при изменении его угла атаки.—*Прим. перев.*

образом, изображения входа и выхода системы оказываются связанными соотношением [1])

$$Y = F_2(s)[F_1(s)X + F_3(s)Y + F_5(s)F_4(s)Y]$$

или

$$\frac{Y}{X} = F_s(s) = \frac{F_1(s)F_2(s)}{1 - F_2(s)[F_3(s) + F_5(s)F_4(s)]}. \qquad (9.85)$$

Следовательно, мы можем изменять передаточную функцию $F_s(s)$ всей системы в целом путем надлежащего формирования передаточной функции $F_4(s)$ цепи регулирования.

Пусть $\Phi(\omega)$ есть спектральная плотность входа x, $g(\omega)$ — спектральная плотность выхода, т. е. перемещения $y(t)$ крыла. Тогда, в силу (9.71),

$$g(\omega) = \Phi(\omega)F_s(i\omega)F_s(-i\omega), \qquad (9.86)$$

где функция $F_s(s)$ определена соотношением (9.85). Вполне очевидно, что для обеспечения наибольшего удобства для пассажиров самолета необходимо уменьшить в наибольшей возможной степени ускорение $y''(t)$ крыла. Это значит, что необходимо достичь минимума величины $\overline{(y'')^2}$. Соотношение (9.16) показывает, что среднее квадратичное значение величины y'' можно подсчитать с помощью функции корреляции. Но функцию корреляции можно найти, коль скоро известна спектральная плотность. Рассматривая эти соотношения совместно с передаточной функцией (9.85) и выражением (9.86) для спектральной плотности выхода, получаем

$$\overline{[y''(t)]^2} = \int_0^\infty \omega^4\Phi(\omega)\left|\frac{F_1(i\omega)F_2(i\omega)}{1 - F_2(i\omega)[F_3(i\omega) + F_5(i\omega)F_4(i\omega)]}\right|^2 d\omega. \qquad (9.87)$$

Задача о нахождении минимума величины $\overline{(y'')^2}$ решается путем построения такой передаточной функции $F_4(s)$ цепи регулирования, которой отвечает искомый минимум. Для этого зададим общий вид передаточной функции этой цепи, оставив параметры пока неопределенными. Тогда при

[1]) В предположении справедливости принципа суперпозиции (поток, создаваемый крылом со щитком, равен сумме потоков от крыла и щитка порознь). — *Прим. перев.*

фиксированных остальных передаточных функциях $F_1(s)$, $F_2(s)$, $F_3(s)$ и $F_5(s)$ и при спектральной плотности входа, заданной, например, с помощью соотношения (9.79), можно подсчитать $\overline{[y''(t)]^2}$ в функции этих неопределенных параметров. В этих условиях задача отыскания минимума сводится к обычной экстремальной задаче по отношению к совокупности этих параметров, которые определяются из условия достижения минимума величины $\overline{(y'')^2}$. Передаточная функция, которую мы получим в результате решения этой экстремальной задачи, окажется наилучшей передаточной функцией в смысле обеспечения наибольших удобств пассажиров в самолете[1]).

Приведенное рассуждение представляет собой лишь один из примеров того, как следует находить оптимальные условия регулирования исходя из поставленной цели и при заданных свойствах входа. При другом способе проектирования системы можно было бы исходить из минимума среднего квадратичного значения упругих напряжений, возникающих в материале крыла под действием турбулентного потока воздуха. В такой постановке задачи оптимальная передаточная функция цепи регулирования будет уже другой, но общая формулировка задачи останется неизменной. Такую задачу о проектировании системы автоматического регулирования при условии достижения оптимальных процессов, определяемых количественными соотношениями, следует рассматривать как дальнейшее продвижение по сравнению с критерием, обеспечивающим только устойчивость системы, и другими качественными критериями при проектировании таких систем, разобранными в предыдущих главах. Общие соображения такого рода впервые были сформулированы, повидимому, А. С. Боксенбомом и Д. Новиком[2]).

---

[1]) Здесь оптимальная передаточная функция находится в рамках заранее предписанной структуры цепи регулирующей обратной связи. Не исключено, что для существенного повышения качества процесса (в данном случае для уменьшения ускорений) придется изменить самую структуру цепи, например путем введения корректирующих контуров и т. п. Эту более общую задачу изложенный метод непосредственно не решает. — *Прим. перев.*

[2]) Boksenbom A. S., Novik D., NACA TN 2939 (1953).

Глава X

РЕЛЕЙНЫЕ СЛЕДЯЩИЕ СИСТЕМЫ [1])

Если следящая система содержит звено, работающее по характеристике «включено — выключено», то такая система называется релейной. Как было отмечено в п. 6.3, одно из больших достоинств релейных систем заключается в их дешевизне. Однако поскольку выход релейного звена изменяется не пропорционально входу, т. е. соотношение между входом и выходом не является линейным, поведение релейных систем нельзя исследовать методами линейной теории. В настоящей главе мы сперва изложим приближенную теорию устойчивости релейных и других подобных следящих систем, основанную на критерии Найквиста в несколько измененной форме. В последующей части этой главы мы рассмотрим более новую и вместе с тем более трудную задачу о методике использования присущих релейным явлениям нелинейных характеристик для улучшения качества следящих систем. К сожалению, этот вопрос исследован пока далеко не полностью, и общее решение его еще впереди.

10.1. Приближенная частотная характеристика реле. Рассмотрим синусоидальный вход $x(t)$ частоты ω и амплитуды a:

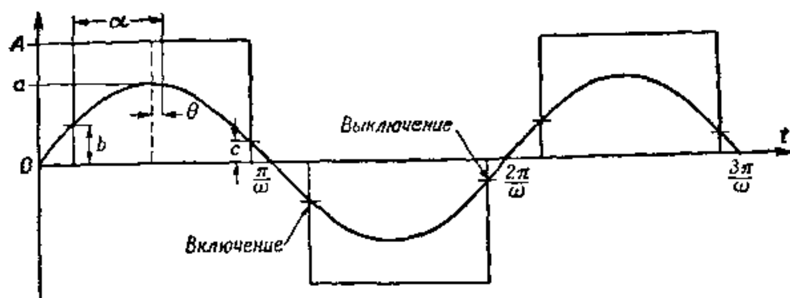
$$x(t) = a \sin \omega t. \tag{10.1}$$

Идеализируем характеристику реле в том смысле, что не будем учитывать продолжительности его срабатывания, считая его действие мгновенным, т. е. без запаздывания во времени. С другой стороны, учтем явление гистерезиса в работе реле: когда входная величина положительна и возрастает, выходная величина реле меняется от нуля до ее наибольшего значения A при $x = b$. Когда входная величина положительна, но убывает, выходная величина

[1]) См. также [5, 6, 7, 9, 13, 16, 22, 42, 76]. — *Прим. перев.*

реле меняется от A до нуля, достигаемого при $x=c$, причем b больше c. Величина b называется током включения, а c — током выключения.

При отрицательных значениях входной величины включение и выключение реле совершаются при $x=-b$ и $x=-c$



Фиг. 71

соответственно. Соотношение между входом и выходом релейного звена можно изобразить с помощью графика, приведенного на фиг. 71. Очевидно, что между входом и выходом имеется сдвиг по фазе. Это отставание по фазе, которое мы обозначим через θ, определяется соотношением

$$\theta=\frac{1}{2}\left[\arcsin\frac{b}{a}-\arcsin\frac{c}{a}\right]. \qquad (10.2)$$

Промежуток времени, в течение которого реле остается включенным, равен $2\alpha/\omega$, где α — половина длины одного импульса прямоугольной волны, определяемая соотношением

$$\alpha=\frac{\pi}{2}+\theta-\arcsin\frac{b}{a}=\frac{1}{2}\left(\pi-\arcsin\frac{b}{a}-\arcsin\frac{c}{a}\right). \qquad (10.3)$$

Период выходной переменной равен периоду входной переменной, т. е. равен $2\pi/\omega$.

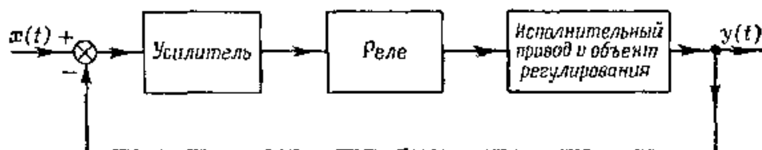
Представим теперь выходную переменную $y(t)$ релейного звена посредством ряда Фурье

$$y(t)=\sum_{1}^{\infty}a_n\sin\left[n(\omega t-\theta)\right]. \qquad (10.4)$$

Для коэффициента a_1 при первой гармонике имеем простое выражение

$$a_1=\frac{4A}{\pi}\sin\alpha,$$

где α определяется с помощью равенства (10.3). В релейной следящей системе, блок-схема которой изображена на фиг. 72, выход релейного звена служит корректирующим сигналом, возбуждающим исполнительный привод системы. Такая следящая система обладает в общем случае свойствами фильтра и в большой мере уменьшает влияние высших гармоник. Поэтому в приближенном



Фиг. 72

исследовании мы можем пренебречь всеми высшими гармониками и принять за выход функцию $a_1 \sin(\omega t - \theta)$. В этом предположении отношение выхода ко входу равно

$$F_r(i\omega) = \frac{4A \sin \alpha}{\pi a} e^{-i\theta}. \tag{10.5}$$

Следовательно, функцию $F_r(i\omega)$ можно рассматривать как частотную характеристику реле. Разумеется, она не представляет собой подлинной частотной характеристики. Подлинная частотная характеристика, определенная таким образом, как это сделано в предыдущих главах, зависит только от частоты и не зависит от амплитуды входа. Однако частотная характеристика, выписанная в правой части равенства (10.5), зависит от амплитуды a входа. Для отрицательных значений ω имеем

$$x(t) = a \sin \omega t = -a \sin |\omega| t, \quad \omega < 0.$$

Следовательно, для такого входа выход может быть приближенно выражен в виде

$$y(t) \approx -\frac{4A}{\pi} \sin \alpha \sin (|\omega| t - \theta) = \frac{4A}{\pi} \sin \alpha \sin (\omega t + \theta), \quad \omega < 0.$$

Тогда отношение выхода к входу

$$F_r(i\omega) = \frac{4A \sin \alpha}{\pi a} e^{i\theta}, \quad \omega < 0.$$

Отсюда, сравнивая с равенством (10.5), имеем

$$F_r(-i\omega) = \overline{F_r(i\omega)}, \quad \omega < 0,$$

и общее свойство частотной характеристики системы, выражаемое равенством (3.17), здесь сохраняется. С другой стороны, функция $F_r(i\omega)$ вообще не зависит от частоты ω. Поэтому использование здесь обозначения частотной характеристики является неподходящим; однако, чтобы подчеркнуть, что эта функция представляет собой отношение выхода ко входу, мы будем пользоваться данным обозначением, несмотря на указанное обстоятельство.

Когда амплитуда a входа очень велика, мы получим, в силу равенств (10.2), (10.3) и (10.5),

$$F_r(i\omega) = \frac{4A}{\pi}\frac{1}{a} \text{ при } a \to \infty. \tag{10.6}$$

Когда амплитуда a очень мала, реле вовсе не сработает. Пороговое значение наступает при $a = b$. Тогда

$$\theta = \alpha = \frac{1}{2}\left(\frac{\pi}{2} - \arcsin\frac{c}{b}\right) \text{ при } a = b. \tag{10.7}$$

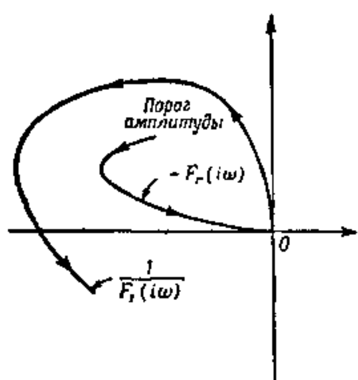
Таким образом, граничные значения частотной характеристики реле полностью определяются характеристикой реле.

10.2. Метод Кохенбургера. Допустим на мгновение, что амплитуды всех гармоник входной функции реле равны между собой. Тогда частотная характеристика $F_r(i\omega)$ реле сводится к комплексной постоянной, определяемой равенством (10.5). Пусть $F_1(i\omega)$ есть частотная характеристика всей остальной части цепи, изображенной на фиг. 72; тогда частотная характеристика всей прямой цепи равна $F_r(i\omega) F_1(i\omega)$. Согласно критерию Найквиста (п. 4.3), для того чтобы система была устойчивой, кривая $1/[F_r(i\omega) F_1(i\omega)]$ на комплексной плоскости при изменении ω от 0 до ∞ должна охватывать точку -1. Другими словами, кривая $1/[F_r(i\omega) F_1(i\omega)]$ должна пересекать вещественную ось слева от точки -1. Но, коль скоро амплитуда a входной функции реле постоянна, будет постоянной и функция $F_r(i\omega)$, и указанное выше условие устойчивости оказывается равносильным требованию о том, чтобы частотная характеристика $1/F_1(i\omega)$ охватывала точку $-F_r(i\omega)$ при изменении ω от 0 до ∞. Такова основа метода Кохенбургера[1]) применения частотных

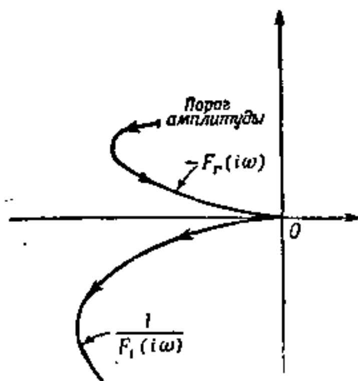
[1]) Kochenburger R. J., Trans. Amer. Inst. Electr. Eng. (AIEE), **69**, 270—284 (1950). (Этот метод был ранее разработан Л. С. Гольдфарбом, см. «Автоматика и Телемеханика», **5**, 349—383 (1947).—*Прим. перев.*)

характеристик для исследования устойчивости релейных следящих систем. Независимо от него аналогичный метод был разработан Дютилем[1]).

Кохенбургер рассуждает следующим образом: если амплитуды гармоник входной функции не равны между собой, то для использования условия устойчивости, описанного в предыдущем пункте, необходимо только применить это условие ко всем значениям $F_r(i\omega)$, соответствующим изменению амплитуды от 0 до ∞. Геометрическое место точек $-F_r(i\omega)$ при различных значениях a представляет собой кривую, которая выходит из точки, отвечающей



Фиг. 73

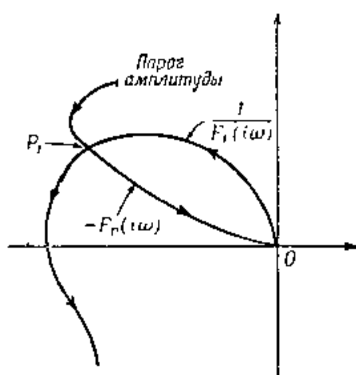


Фиг. 74

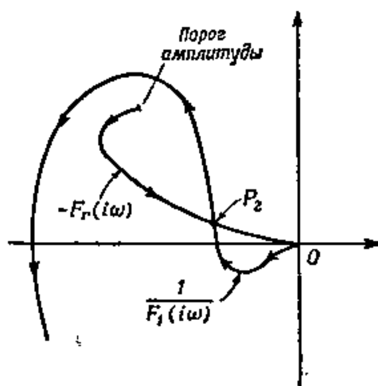
пороговому значению амплитуды, определяемому соотношением (10.7), и оканчивается в начальной точке 0 плоскости комплексной переменной (фиг. 73). Условие Кохенбургера, достаточное для устойчивости исследуемой системы, формулируется следующим образом: кривая $1/F_1(i\omega)$ должна охватывать геометрическое место $-F_r(i\omega)$ целиком, как это показано на фиг. 73. На фиг. 74 показано течение соответствующих кривых в случае полной неустойчивости системы. Стрелки, нанесенные на кривой $-F_r(i\omega)$, указывают направление возрастания амплитуд гармоник входной функции реле. Стрелки, нанесенные на кривой $1/F_1(i\omega)$, указывают направление возрастания частоты (частота возрастает начиная со значения $\omega = 0$ в начале).

[1]) Dutilh J. R., L'Onde électrique, 30, 438—445 (1950).

Помимо этих случаев полной устойчивости и полной неустойчивости системы, встречаются различные случаи частичной устойчивости или неустойчивости с возможностью существования автоколебаний определенной частоты и постоянной амплитуды. Например, на фиг. 75 изображен ход кривых, соответствующих устойчивым автоколебаниям. При малых амплитудах точки кривой $-F_r(i\omega)$ находятся «вне» кривой $1/F_1(i\omega)$ и, следовательно, система неустойчива, ибо амплитуда колебаний будет нарастать. С увеличением амплитуды точка $-F_r(i\omega)$ перемещается по направлению к кривой $1/F_1(i\omega)$ и, в конце концов, достигает точки P_1 пересечения



Фиг. 75

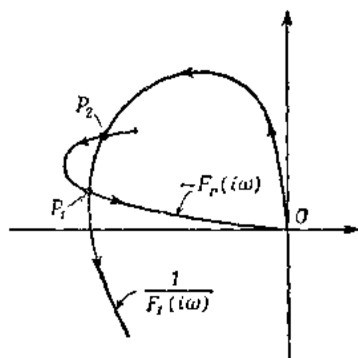


Фиг. 76

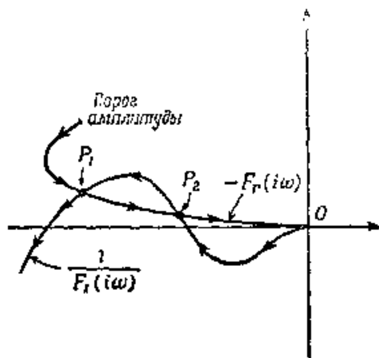
кривых, в которой система будет колебаться с частотой и амплитудой, соответствующими этой точке. Таким образом, это колебание начинается при достаточно малых начальных отклонениях, и все явление получило название мягкого возбуждения. Система не стремится удалиться от состояния, отвечающего точке P_1, так как всякому увеличению амплитуды будет противодействовать демпфирующее действие при попадании изображающей точки в область устойчивости. Таким образом, точка P_1 характеризует сходящиеся колебания, и в ней система будет всегда находиться в состоянии автоколебаний. На фиг. 76 изображен другой случай, когда точка P_2 пересечения кривых $-F_r(i\omega)$ и $1/F_1(i\omega)$ соответствует неустойчивым колебаниям. Система всегда будет стремиться удалиться от состояния, отвечаю-

щего точке P_2. Однако при достаточно малых возмущениях она остается устойчивой.

Кривые на фиг. 77 и 78 отвечают более сложным случаям, при которых одновременно имеются как точки типа P_1, отвечающие устойчивым автоколебаниям, так и точки типа P_2, отвечающие неустойчивым автоколебаниям. Система, соответствующая фиг. 77, будет колебаться при сообщении ей возмущений достаточно большой амплитуды. Такое явление называется жестким возбуждением. Система, соответствующая фиг. 78, всегда будет колебаться, если только амплитуда возмущения не слишком велика. При достаточно больших возмущениях колебания будут расходящимися.



Фиг. 77



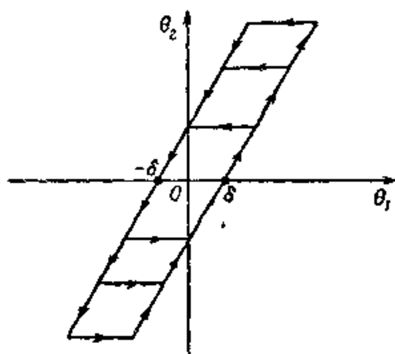
Фиг. 78

Во всех случаях, соответствующих фиг. 75—78, наблюдается характерная зависимость поведения системы от амплитуды возбуждающего ее воздействия и возможность существования установившихся колебаний при определенных как частоте, так и амплитуде. Все эти свойства характерны для нелинейных систем, и ими не обладают системы линейные, изучению которых были посвящены предыдущие главы[1]). Конечно, существования таких свойств можно было ожидать уже на основании введения в теорию нелинейных систем, содержащегося в п. 1.3.

---

[1]) Следует подчеркнуть, что изложенный выше метод является приближенным и, будучи весьма практичным, в некоторых случаях может привести даже к неправильным выводам.— *Прим. ред.*

10.3. Другие нелинейные устройства, не пропускающие высоких частот. Метод Кохенбургера представляет собой весьма эффективный метод решения задачи об устойчивости релейной следящей системы. Его можно применять при исследовании достаточно сложных систем; при этом метод позволяет непосредственно использовать частотные характеристики, полученные экспериментальным путем. В большинстве прикладных задач, где исполнительный привод, следующий за реле в схеме системы, обладает достаточными фильтрационными свойствами, пренебрежение

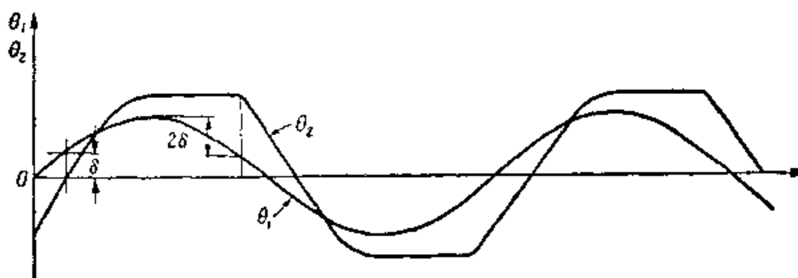


Фиг. 79

высшими гармониками является вполне оправданным. Свойства системы, найденные теоретически, совпали со свойствами, обнаруженными в результате эксперимента. Поэтому, коль скоро при разработке системы ставится только требование обеспечения устойчивости релейной системы, метод Кохенбургера полностью решает задачу об обеспечении устойчивости такой системы.

Вместе с тем метод Кохенбургера применим не только к релейным следящим системам, но в такой же мере ко многим другим нелинейным устройствам. Отправным пунктом этого метода исследования служит то обстоятельство, что качественный характер отработки реле не зависит от частоты, но зависит от амплитуды, тогда как качественный характер отработки линейной системы зависит от частоты и не зависит от амплитуды.

Но, как известно, имеется много нелинейных устройств, процессы в которых протекают в качественном отношении так же, как и в релейных звеньях. Примером такого устройства служит механическая передача с зазором, в чем можно убедиться с помощью следующего рассуждения: пусть θ_1 — угловая координата ведущего вала, на который жестко насажена входная шестерня передачи, и пусть θ_2 — угловая координата ведомого вала и жестко связанной с ним выходной шестерни передачи. Тогда соотношение между θ_1 и θ_2 можно изобразить с помощью графиков, подобных графику, приведенному на фиг. 79, где 2δ — полная величина зазора в передаче. Если вход θ_1 передачи изменяется по синусоидальному закону, то ее выход изображается определенным образом сплюснутой синусообразной волной,



Фиг. 80

отстающей по фазе от входа (фиг. 80). Так как соотношение, связывающее θ_1 и θ_2, не зависит явно от времени, то легко убедиться, что форма выходного колебания θ_2 не изменится при изменении частоты входного колебания θ_1. Таким образом, реакция механической передачи изменяется по амплитуде, но остается неизменной по частоте.

Обозначим отношение амплитуды основной гармоники колебания θ_2 к амплитуде основной гармоники колебания θ_1 через $F_g(i\omega)$, причем $F_g(i\omega)$ характеризует также и фазовый сдвиг; тогда последнее утверждение означает, что $F_g(i\omega)$ является функцией от a, но не зависит от ω. Следовательно, эту зависимость можно применять для изучения следящих систем, содержащих механические передачи с зазорами, совершенно так же, как зависимость $F_r(i\omega)$ релейного звена использовалась в предыдущем разделе.

10.4. Оптимальная отработка релейной следящей системы. Напомним, что, согласно критерию устойчивости Кохенбургера, кривая $1/F_1(i\omega)$ должна охватывать всю кривую $-F_r(i\omega)$ целиком, а не только точку -1, как в случае обычных следящих систем. Это условие накладывает более сильные ограничения на другие звенья системы; именно по этой причине качество релейной следящей системы в общем случае ниже качества обычной следящей системы. Однако это свойство не представляет собой недостатка, присущего релейным следящим системам. В самом деле, вместо того, чтобы ограничиваться обеспечением устойчивости, мы можем рассматривать реле как устройство, предназначенное для операций включения и выключения и способное производить положительный или отрицательный постоянный корректирующий сигнал или не вырабатывать вовсе никакого сигнала, и затем поставить вопрос: каким образом следует производить включение и выключение реле в зависимости от выходной переменной, с тем чтобы получить оптимальный процесс во всей системе в целом. Пусть, например, требуется обеспечить возможно более быстрое возвращение системы к стабилизируемому положению после возмущения. Такое требование не только предусматривает возвращение системы к стабилизируемому состоянию (т. е. устойчивость ее), но также обусловливает быстрейшее осуществление этого возвращения. Решение данной оптимальной задачи заключается в указании закона срабатывания реле в зависимости от изменения выходной переменной; знание этого закона срабатывания лежит в основе расчета следящей системы.

Релейные следящие системы, разработанные с учетом требований, вытекающих из рассмотренной здесь более общей постановки задачи, будут, несомненно, превосходить по качеству обычные линейные следящие системы по той причине, что нелинейная характеристика релейных звеньев используется при этом самым полным образом[1]).

10.5. Фазовая плоскость. Обозначим через y выходную координату, а через x — входную; тогда дифференциальное

---

[1]) Мало того, доказано, что именно системы, содержащие релейные звенья, являются *наиболее* быстродействующими из всех возможных. — *Прим. ред.*

уравнение системы второго порядка, линейной или нелинейной, можно в общем виде написать следующим образом:

$$f(y,\ \dot{y},\ \ddot{y};\ t) = x(t), \tag{10.8}$$

где точка означает дифференцирование по времени. Уравнение (10.8) можно переписать в виде системы уравнений:

$$\left.\begin{aligned} f\left(y,\ \dot{y},\ \frac{d\dot{y}}{dt};\ t\right) &= x(t), \\ \frac{dy}{dt} &= \dot{y}. \end{aligned}\right\} \tag{10.9}$$

Если рассматривать y и $\dot{y}$ как зависимые переменные, то уравнения (10.9) представляют собой систему двух дифференциальных уравнений первого порядка относительно неизвестных y и $\dot{y}$. Если входное воздействие отсутствует ($x = 0$) и если функция f не зависит явно от t, т. е. система является автономной, что часто и имеет место, то первое уравнение системы (10.9) можно разрешить[1]) относительно $d\dot{y}/dt$ и выразить $d\dot{y}/dt$ как функцию от y и $\dot{y}$. После этого систему можно переписать в виде

$$\left.\begin{aligned} \frac{d\dot{y}}{dt} &= \dot{y}k(y,\ \dot{y}), \\ \frac{dy}{dt} &= \dot{y}. \end{aligned}\right\} \tag{10.10}$$

Эта система уравнений не содержит t явно. Разделив первое уравнение системы (10.10) на второе, получим

$$\frac{d\dot{y}}{dy} = k(y,\ \dot{y})\,^{2)}. \tag{10.11}$$

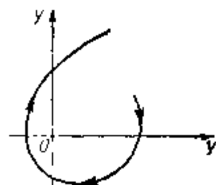
---

[1]) Всюду, где $\dfrac{\partial f}{\partial \ddot{y}} \neq 0$. — *Прим. перев.*

[2]) Правая часть первого уравнения (10.10) обозначается через $\dot{y}k(y,\ \dot{y})$ для удобства, чтобы записать производную $d\dot{y}/dy$ в виде функции от y и $\dot{y}$. — *Прим. перев.*

Мы пришли, таким образом, к уравнению первого порядка, в котором y играет роль независимой, а $\dot{y}$ — зависимой переменной. Решив это уравнение, мы можем с помощью системы (10.10) определить зависимость между t и y.

С физической точки зрения только что проведенные преобразования предполагают, что состояние системы описывается с помощью переменных y и $\dot{y}$, а не с помощью более принятых для этой цели переменных y и t. Если переменная y определяет положение некоторой материальной точки, то $\dot{y}$ представляет собой ее скорость. Таким образом, переменную $\dot{y}$ можно рассматривать как величину, характеризующую количество движения этой материальной точки. Тогда y и $\dot{y}$ характеризуют положение и количество движения материальной точки соответственно. Физики называют такое описание состояния системы изображением в фазовом пространстве. В рассмотренном частном примере фазовое пространство имеет лишь два измерения и, таким образом, представляет собой фазовую плоскость. Следовательно, поведение системы второго порядка определяется кривой на фазовой плоскости. Каждая точка этой кривой изображает состояние системы в определенный момент t. Направление на кривой, соответствующее возрастанию времени, принято указывать стрелками вдоль этой кривой (фиг. 81). Если система имеет порядок n, то фазовое пространство будет пространством n измерений, а поведение системы будет изображаться кривой в этом n-мерном пространстве.



Фиг. 81

Практическое достоинство метода изображения на фазовой плоскости заключается в том, что очень большое количество нелинейных систем второго порядка принадлежит к классу автономных систем и для них справедливо уравнение (10.11), а это уравнение можно решить, по крайней мере графически, по методу изоклин[1]). В самом деле,

---

[1]) См. [1, 20, 52, 53]. — *Прим. перев.*

характер системы сразу же выявляется, как только на фазовой плоскости наносится поле направлений фазовых траекторий в соответствии с уравнением (10.11). На использовании таких геометрических свойств фазовой плоскости зиждется теория нелинейных колебаний; метод, связанный с использованием указанных геометрических свойств, называется топологическим методом нелинейной механики[1]).

Чтобы выразить наши предыдущие понятия на языке изображений на фазовой плоскости, рассмотрим простую задачу об исследовании линейной системы второго порядка, свободной от действия возмущающей силы,

$$\frac{d^2y}{dt^2} + 2\zeta \frac{dy}{dt} + y = 0. \qquad (10.12)$$

Это уравнение получится с помощью уравнения (3.39), если в этом последнем уравнении выбрать единицу времени так, чтобы сделать частоту ω_0 свободных незатухающих колебаний равной единице. Коэффициент ζ представляет собой, конечно, отношение коэффициента демпфирования системы к его критическому значению. В колебательных системах $|\zeta| < 1$[2]) уравнение (10.12) можно переписать в виде

$$\left.\begin{aligned} \frac{d\dot{y}}{dt} &= -2\zeta\dot{y} - y, \\ \frac{dy}{dt} &= \dot{y}. \end{aligned}\right\} \qquad (10.13)$$

Тогда уравнение, соответствующее (10.11), будет иметь вид

$$\frac{d\dot{y}}{dy} = -\frac{2\zeta\dot{y} + y}{\dot{y}} = -2\zeta - \frac{y}{\dot{y}}. \qquad (10.14)$$

Очевидно, что линиями постоянного наклона $d\dot{y}/dy$ на фазовой плоскости служат радиальные прямые, проходящие через начало координат. На фиг. 82—86 изображены пять

---

[1]) См. [1, 8, 53]. — *Прим. перев.*

[2]) Следует иметь в виду большое качественное различие в поведении систем, в которых $0 < \zeta < 1$, и систем, в которых $-1 < \zeta < 0$. — *Прим. перев.*

типов фазовых траекторий, отвечающих случаям $\zeta < -1$; $-1 < \zeta < 0$; $\zeta = 0$; $0 < \zeta < 1$ и $1 < \zeta$ соответственно. Графики на фиг. 82 и 86 характеризуют неколебательные системы, графики на фиг. 83 и 85 соответствуют колебательным системам, а график на фиг. 84 относится к случаю чисто гармонических колебаний.

На этих фигурах начало координат на фазовой плоскости соответствует состоянию равновесия, при котором обе производные dy/dt и $d\dot{y}/dt$ равны нулю. В математическом отношении начало координат представляет собой особую точку системы уравнений (10.13). Однако в каждом из трех различных случаев $\zeta < 0$, $\zeta = 0$ и $0 < \zeta$ состояние равновесия имеет различный характер. Из графиков на фиг. 82 и 83 следует, что при $\zeta < 0$ все фазовые траектории расходятся от состояния равновесия. Поэтому начало координат является точкой неустойчивого равновесия. Из графиков на фиг. 85 и 86 следует, что при $0 < \zeta$ все фазовые траектории сходятся к состоянию равновесия. Поэтому начало координат является точкой устойчивого равновесия. Математически начало координат на фиг. 82 и 86 представляет собой точку, через которую проходят все фазовые траектории; эта точка называется узлом. Начало координат на фиг. 83 и 85 служит центром спиралей и называется фокусом. В особом случае, отвечающем фиг. 84, где $\zeta = 0$, фазовые траектории окружают начало, и потому в этом случае начало называется центром.

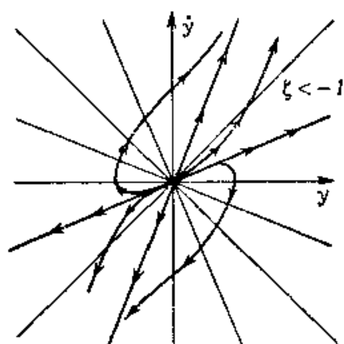
Если основное уравнение, описывающее систему второго порядка, содержит постоянный возмущающий член, т. е. имеет вид

$$\frac{d^2y}{dt^2} + 2\zeta\frac{dy}{dt} + y = c, \tag{10.15}$$

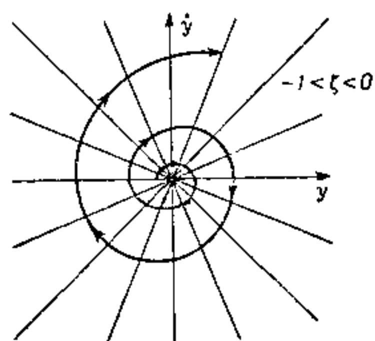
где c — постоянная, то его можно записать следующим образом:

$$\frac{d^2(y-c)}{dt^2} + 2\zeta\frac{d(y-c)}{dt} + (y-c) = 0.$$

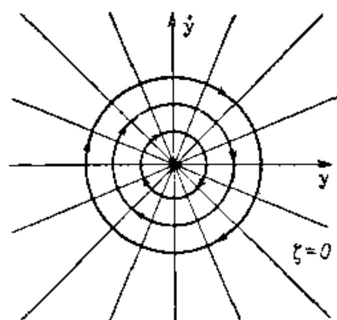
Следовательно, в этом случае течение фазовых траекторий совершенно аналогично их течению, изображенному на фиг. 82 и 86, с тем единственным изменением, что в рассматриваемом случае точка равновесия сдвинута по оси y и находится в точке $y = c$ на этой оси.



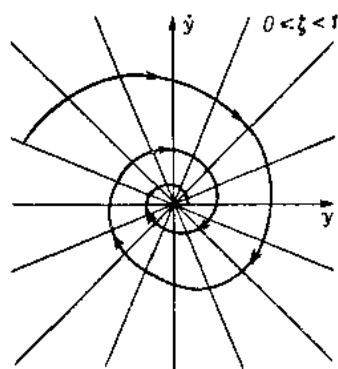
Фиг. 82



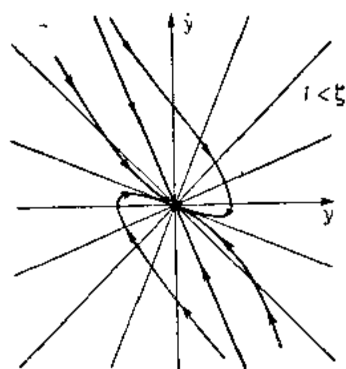
Фиг. 83



Фиг. 84



Фиг. 85



Фиг. 86

10.6. Линейное включение. В последующем рассуждении мы упростим задачу о переключении реле, предположив, что реле может находиться лишь в одном из двух состояний: в одном из них выходная координата равна положительной, а в другом — отрицательной единице. То обстоятельство, что абсолютная величина выходной координаты реле принята равной единице, очевидно, не ограничивает общности рассуждения. Но прежде, чем переходить к разбору задачи об оптимальном законе срабатывания реле в общем виде, рассмотрим более простой случай линейного закона срабатывания: возмущающая функция c, создаваемая релейным устройством, равна единице по абсолютной величине, а знак ее совпадает со знаком линейной формы $ay+b\dot{y}$. Мы касаемся этого случая в целях иллюстрации некоторых свойств общей задачи.

Свойства релейных следящих систем с линейным законом срабатывания исследовали Флюгге-Лотц[1]) и Флюгге-Лотц и Клоттер[2]). Ниже приводится обзор качественных результатов их исследований. Упомянутые авторы рассматривали систему, описываемую дифференциальным уравнением

$$\frac{d^2y}{dt^2}+2\zeta\frac{dy}{dt}+y=\operatorname{sign}(ay+b\dot{y}), \qquad 0<\zeta<1. \quad (10.16)$$

Фазовую траекторию, соответствующую положительной единичной вынуждающей функции, назовем P-дугой, а фазовую траекторию, соответствующую отрицательной единичной вынуждающей функции, — N-дугой. Совокупность всех P-дуг называется P-совокупностью, а совокупность всех N-дуг — N-совокупностью. Вспоминая наши рассуждения, связанные с исследованием уравнения (10.15), мы приходим к заключению, что для данного уравнения совокупность P представляет собой систему сходящихся спиралей с фокусом в точке $y=+1$, $\dot{y}=0$, а совокупностью N служит система сходящихся спиралей с фокусом в точке

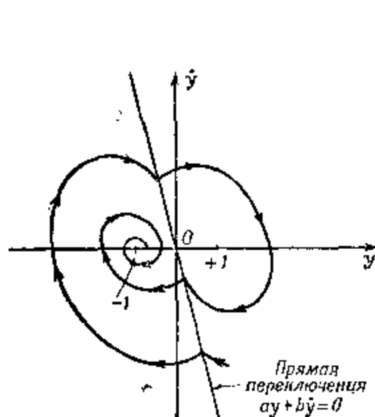
---

[1]) Flügge-Lotz I., Z. angew. Math. Mech., **25—27**, 97—113 (1947).

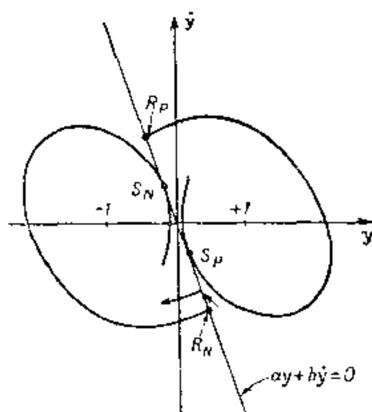
[2]) Flügge-Lotz I., Klotter K., Z. angew. Math. Mech., **28**, 317—337 (1948).

$y = -1$, $\dot{y} = 0$. Желаемое окончательное состояние системы соответствует, конечно, началу координат $y = \dot{y} = 0$.

В зависимости от сочетания знаков постоянных a и b в функции в правой части уравнения (10.16), определяющей закон переключения реле, можно выделить четыре случая. Отнесем к случаю I действие системы при $a > b$; $b > 0$. В этом случае линия переключения реле $ay + b\dot{y} = 0$ представляет собой на фазовой плоскости прямую, проходящую



Фиг. 87



Фиг. 88

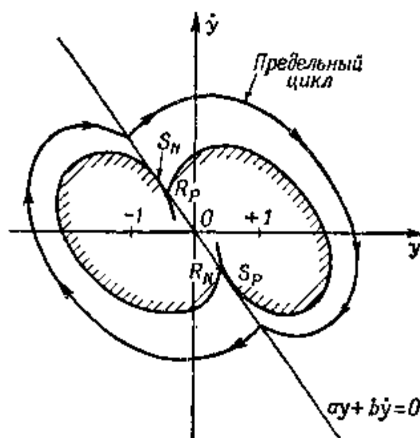
через начало координат и расположенную во втором и четвертом квадрантах. Вправо от этой прямой функция $ay + b\dot{y}$ положительна, и соответствующая область служит областью расположения P-совокупности. Влево от прямой функция $ay + b\dot{y}$ отрицательна, и соответствующая область служит областью расположения N-совокупности. На прямой переключения обе совокупности смыкаются, образуя угловые точки (фиг. 87). Условие существования периодического движения состоит в том, чтобы имелась такая P-дуга, обе точки пересечения которой с прямой переключения находятся на одинаковом расстоянии от начала координат; тогда, в силу симметрии, будет иметься и такая N-дуга с другой стороны прямой переключения, которая соединяет

эти точки пересечения. Обе эти дуги, рассматриваемые совместно, образуют на фазовой плоскости замкнутую кривую. По терминологии нелинейной механики такие периодические движения называются предельными циклами. Мы сейчас убедимся, что в нашей задаче периодическое движение может иметь место.

Условимся обозначать символами S_P и S_N те точки на прямой переключения, в которых P-дуга и N-дуга соответственно касаются этой прямой, а символами R_P и R_N—последние перед S_P и S_N точки пересечения дуг P и N с этой же прямой. Точки S_P и S_N расположены симметрично относительно начала координат, так же как и точки R_P и R_N. Пусть числа ζ, a и b имеют такие значения, что точка R_N находится вне отрезка $S_P S_N$; такое расположение этих точек показано на фиг. 88. Всякое решение, начинающееся в достаточной близости от отрезка $S_P S_N$, будет удаляться от прямой переключения в том или другом направлении, смотря по тому, с какой стороны от прямой переключения находится начальная точка решения, и никогда уже не приблизится снова к прямой переключения. Можно также показать, что конечная точка каждой P-дуги, лежащей вправо от прямой переключения и оканчивающейся на этой прямой, расположена ближе к началу координат, чем ее начальная точка; отсюда следует, что в данном предположении условие периодичности движения никогда не может быть выполнено. Следовательно, фазовые траектории системы всегда будут иметь форму спирали, наматывающейся на один из фокусов, а конечное состояние системы не будет совпадать с началом координат на фазовой плоскости.

Если же, однако, точки R_N и R_P лежат внутри отрезка $S_P S_N$ (фиг. 89), то периодическое движение существует. Причина этого заключается в следующем: P-дуга $R_P S_P$, выходит из точки, которая расположена, по предположению, ближе к началу координат, чем ее конечная точка. Но на большом удалении от начала, где влиянием сдвига фокусов семейств сходящихся спиралей в точки $+1$ и -1 для P-совокупности и N-совокупности соответственно можно пренебречь, всякая P-дуга должна начинаться на прямой переключения в точке, расположенной от начала координат дальше, чем конечная точка этой P-дуги, также лежащая на прямой переключения. Следовательно, такие

дуги обладают свойствами, обратными по отношению к свойствам дуг $R_P S_P$. Поэтому в силу непрерывности некоторая промежуточная P-дуга этого типа должна начинаться и оканчиваться в точках, одинаково удаленных от начала координат. Это свойство и представляет собой условие существования периодического решения, изображенного на фиг. 89 в виде предельного цикла. Как показывает более подробное исследование, это решение действительно



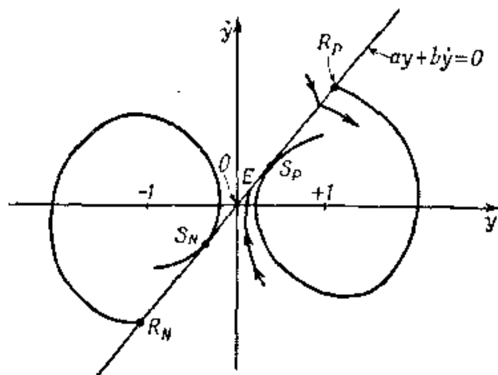
Фиг. 89

существует и является единственным, а также, что еще важнее, это периодическое решение орбитально устойчиво; все рёшения, начинающиеся вне предельного цикла, наматываются на него, и подобным же образом наматываются на предельный цикл решения, начинающиеся внутри этого цикла, но вне области, ограниченной контуром $R_P S_P R_N S_N$ и заштрихованной на фиг. 89[1]). Фазовые траектории, начинающиеся внутри заштрихованной области, наматываются на один из фокусов. В данном случае, как и выше, изображающая точка не может достигнуть начала координат за конечный промежуток времени.

---

[1]) Такой предельный цикл называется устойчивым.—*Прим. ред.*

Перейдем к случаю II: $a>0$, $b<0$. Прямая переключения $ay+b\dot{y}=0$ находится теперь в первом и третьем квадрантах. При этом P-совокупность и N-совокупность оказываются теми же самыми, как и в первом случае. В случае II периодическое решение не существует. Пусть символами R_P, R_N, S_P и S_N обозначены такие же точки, как и в первом случае; тогда точки R_P, S_P, O, S_N и R_N будут лежать на прямой переключения в порядке, указанном на фиг. 90.



Фиг. 90

Но в интервале S_PS_N возникает новое явление. Рассмотрим какое-либо решение, достигающее прямой переключения в этом интервале, например в точке E. Каково будет дальнейшее течение фазовой траектории, достигшей этой точки? Так как переключение реле производится именно на этой прямой, изображающая точка должна перемещаться вдоль некоторой N-дуги. Однако N-дуга уходит из точки E обратно в ту же полуплоскость, из которой точка, изображающая рассматриваемое решение, пришла в точку E, а по эту сторону от прямой переключения решение может изображаться только P-дугами. С другой стороны, изображающая точка, безусловно, не может перемещаться вдоль P-дуги, выходящей из точки E и расположенной по другую сторону от прямой переключения. Следовательно, мы должны сказать, что в моменты времени, следующие за моментом достижения изображающей точкой точки E, решение не определено; оно «оканчивается» в точке E. Всякое решение, выходящее из точек вне интервала R_PR_N, приближается

по дуге спирали по направлению к началу координат до тех пор, пока характеризующая его фазовая траектория не достигнет прямой переключения в интервале $R_N R_P$. Если же она достигает этой прямой в интервале $R_P S_P$ или $R_N S_N$, она будет наматываться соответственно на фокус $+1$ или -1. Если же фазовая траектория достигает прямой переключения в интервале $S_P S_N$, то наблюдается любопытное явление «остановки» движения.

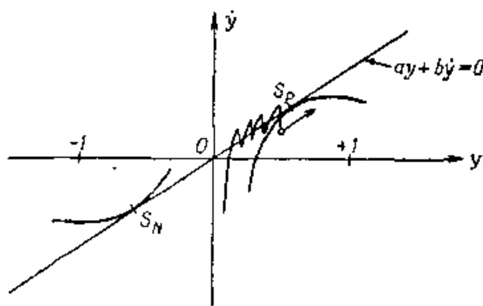
В действительности процесс в системе не может «прекратиться», а должен продолжаться[1]). Для разрешения этого парадокса обратим внимание на то обстоятельство, что процесс переключения всегда сопряжен с наличием некоторого запаздывания по времени. Когда фазовая траектория, характеризующая движение, достигает прямой переключения, то в реальных условиях эта траектория пересекает прямую переключения и заходит за нее на небольшое расстояние прежде, чем происходит изменение знака возмущающей функции. В случае I такое запаздывание (если только оно не слишком велико) не отражается существенным образом на поведении системы. Но в данном случае благодаря инерции системы в ней не наступает необходимого «прекращения» движения. Рассмотрим движение, достигающее такой конечной точки: из-за запаздывания в переключении реле это движение уже не окончится в этой точке, а будет продолжаться, так что характеризующая его фазовая траектория несколько перейдет за эту точку; после этого произойдет переключение реле и на фазовой траектории появится угловая точка, соответствующая моменту переключения, причем в этой точке решение остается определенным. Исходя из этой угловой точки фазовая траектория пересечет прямую переключения в обратном направлении, зайдет за нее на небольшое расстояние, после чего вновь образуется угловая точка и т. д., как изображено на фиг. 91. Из рассмотрения этой фигуры видно также, что такое зигзагообразное действие системы сводится к сползанию изображающей точки из зоны, прилегающей к интервалу $S_P S_N$, по направлению к одному из фокусов, на который фазовая траектория будет, в конце

---

[1]) Так как в точке E скорость $\dot{y}>0$ и, следовательно, точка E не может быть точкой равновесия (покоя). Точки покоя ($y=0$) могут лежать только на оси абсцисс.—*Прим. перев.*

концов, наматываться. Но, хотя благодаря влиянию инерции системы и не наступает прекращения движения, поведение системы все же является неудовлетворительным, поскольку решение не способно достичь начала координат.

Случай III определяется неравенствами $a < 0$, $b > 0$. Здесь прямая переключения опять лежит в первом и третьем квадрантах фазовой плоскости. Но P-совокупность расположена теперь по левую сторону от прямой переключения,



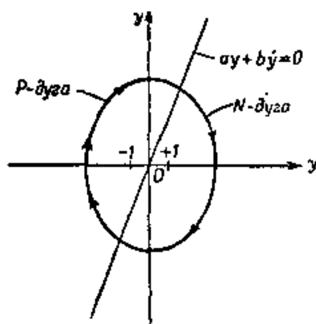
Фиг. 91

а N-совокупность—по правую сторону от нее. В этом случае всегда существует устойчивое периодическое решение, т. е. устойчивый предельный цикл (фиг. 92); этот цикл определяет свойства системы, поскольку все остальные решения на него наматываются. При этом решение опять не может достичь начала координат фазовой плоскости.

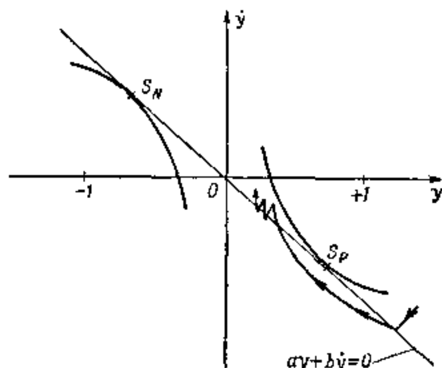
Случаю IV соответствуют отрицательные значения обеих постоянных a и b. Прямая переключения расположена так же, как и в случае I, но течение фазовых траекторий аналогично их течению в случае III. Можно показать, что в случае IV периодические решения не существуют. Без учета запаздывания в срабатывании реле отрезок S_PS_N (фиг. 93) состоит из конечных точек. Проводя фазовые траектории, оканчивающиеся в этих точках, можно убедиться, что эти траектории заполняют всю плоскость. Таким образом, в этом случае все движения «оканчиваются» на отрезке S_PS_N прямой переключения.

Но, как и в случае II, картина движения меняется при учете запаздывания. Это запаздывание не оказывает никакого существенного влияния на движение до тех пор, пока

соответствующая фазовая траектория не достигнет отрезка $S_P S_N$. В этот момент изображающая точка вместо остановки перейдет на небольшое расстояние за прямую переключения, после чего образуется угловая точка фазовой траектории; затем изображающая точка опять пересечет прямую переключения, снова образуется угловая точка фазовой траектории, и т. д. Как видно из рассмотрения фиг. 93, эти колебания направляют систему к началу координат. В конце концов, в системе установятся колебания высокой частоты и малой амплитуды около начала координат; при этом,



Фиг. 92



Фиг. 93

частота колебаний будет тем выше, чем меньше инерция системы. В таком случае говорят, что следящая система работает в вибрационном режиме (или «рыскает»).

Таким образом, из четырех рассмотренных случаев только в одном — четвертом — система обладает свойством стремления к требуемому равновесному состоянию. Но и в этом случае система совершает автоколебания малой амплитуды около состояния равновесия. Изложенное нами исследование Флюгге-Лотца и Клоттера иллюстрирует недостатки, обусловленные линейностью закона переключения. Отсюда следует вывод, что с помощью линейного закона переключения никак нельзя получить оптимального процесса в релейной следящей системе.

10.7. Оптимальный закон включения. Уравнения автономной системы второго порядка в предположении, что

имеется возмущающая функция, абсолютная величина которой равна единице, могут быть написаны в виде

$$\left.\begin{aligned} \frac{dy}{dt} &= \dot{y}, \\ \frac{d\dot{y}}{dt} &+ g(y, \dot{y}) = \varphi(y, \dot{y}). \end{aligned}\right\}^{1)} \qquad (10.17)$$

Здесь $\varphi(y, \dot{y})$ — разрывная функция, способная принимать только два значения: $+1$ или -1. При этих условиях задачу о нахождении оптимального закона включения можно сформулировать следующим образом: требуется разыскать функцию $\varphi(y, \dot{y})$, обладающую тем свойством, что решение, начинающееся в любой точке p фазовой плоскости, будет проходить через начало O координат, а промежуток времени, в течение которого изображающая точка переместится от точки p до начала O, будет минимальным в том смысле, что ни при каком другом выборе функции φ этот промежуток времени не будет короче. Функцию $\varphi(y, \dot{y})$, удовлетворяющую этим условиям, назовем оптимальной функцией включения. Задачу об оптимальном законе включения рассматривал Бушау [2]; он же полностью решил эту задачу[3]) для частного случая, когда функция $g(y, \dot{y})$ линейна, т. е. $g(y, \dot{y}) = 2\zeta\dot{y} + y$ при всех вещественных значениях ζ. Однако математические рассуждения, с помощью которых Бушау получил свои результаты, сложны и их трудно распространить на другие случаи. Поэтому мы ограничимся лишь указанием на методику его решения.

Общий результат, пригодный для любой непрерывной функции $g(y, \dot{y})$, доставляет *каноническая траектория* Бушау (траектория на фазовой плоскости). Так как функ-

---

[1]) Обычно вынуждающей силой (функцией) называют силу (функцию), зависящую явно от независимой переменной t. Здесь это название сохранено для функции φ переключения, не зависящей явно от t. — *Прим. перев.*

[2]) Bushaw D. W., Experimental Towing Tank, Stevens Institute of Technology, Report 469, Hoboken, N. J., 1953.

[3]) См. также [76—78], а также [5], т. II. — *Прим. ред.*

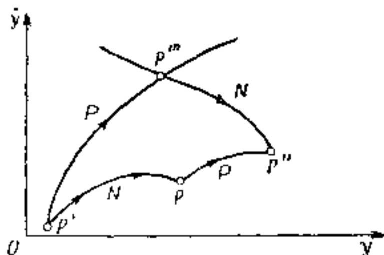
ция $\varphi(y, \dot{y})$ может принимать только значения $+1$ или -1, то такая траектория состоит из P-дуг и N-дуг. Точку соединения этих дуг будем называть угловой точкой типа PN, если с ростом времени значение функции переключения в этой точке меняется от $+1$ к -1. Подобным же образом мы назовем точку соединения этих дуг угловой точкой типа NP, если значение функции φ переключения в этой точки меняется от -1 к $+1$. Траектория называется канонической, если она не содержит угловых точек типа NP над осью y и угловых точек типа PN под осью. Важность понятия канонической траектории объясняется тем, что оптимальная фазовая траектория, т. е. траектория минимального времени прихода системы в начало координат, должна быть траекторией канонической. Это означает следующее: если система, начиная от точки p, будет двигаться вдоль некоторой траектории Δ, не являющейся канонической, то можно найти каноническую траекторию, выходящую из той же точки p, которая, в единицах времени, «короче» траектории Δ. Это можно показать совсем просто. Пусть, например, задана траектория, содержащая угловую точку p типа NP над осью y (фиг. 94). Обозначим через p' ту из указанных ниже двух точек, которая ближе к p, а именно: последнюю угловую точку траектории, предшествующую точке p, или последнюю точку пересечения траектории с осью y, также предшествующую точке p. Через p'' обозначим точку, следующую за точкой p и определенную аналогично точке p'. Определив эти две точки, проведем от точки p P-дугу в направлении возрастания t и N-дугу в направлении убывания t. Тогда, в силу основных уравнений (10.17), получим

$$\frac{d\dot{y}}{dy} = \frac{-g(y, \dot{y}) + \varphi(y, \dot{y})}{\dot{y}}. \tag{10.18}$$

Поэтому в любой точке фазовой плоскости алгебраическое значение наклона P-дуги всегда больше наклона N-дуги [1]).

[1]) Речь идет о тангенсе угла наклона или о величине производной в данной точке для траектории типа P и N. Так как $+1 > -1$ [см. (10.18)], то этот тангенс больше в данной точке для P-дуги, чем для N-дуги. — *Прим. ред.*

Вследствие этого конфигурация траекторий должна иметь вид, изображенный на фиг. 94. Если мы теперь изменим эту траекторию, заменив ее участок $p'pp''$ участком $p'p'''p''$, то тем самым мы удалим угловую точку типа NP и траектория станет канонической. Обозначив через $t(p'pp'')$ продолжительность движения системы от p' до p'' через p и через $t(p'p'''p'')$ продолжительность ее движения по соответствующему участку канонической траектории, напишем



Фиг 94.

$$t(p'pp'') = \int\limits_{p'pp''} \frac{dy}{\dot{y}},$$

$$t(p'p'''p'') = \int\limits_{p'p'''p''} \frac{dy}{\dot{y}}.$$

Но при любых значениях y величина скорости $\dot{y}$ на канонической траектории больше, чем на первоначально взятой траектории, и поэтому

$$t(p'p'''p'') < t(p'pp'').$$

Таким образом, каноническая траектория «короче» неканонической, т. е. требует для прохождения меньшего времени.

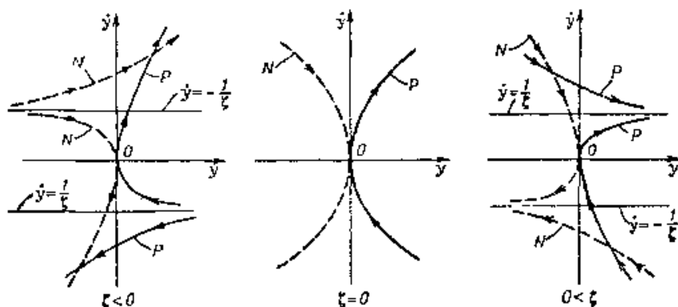
В качестве простого примера применения теории об оптимальной функции переключения рассмотрим систему, в которой $g(y, \dot{y}) = \zeta\dot{y}$.

В этом случае система уравнений (10.17) принимает вид

$$\left.\begin{aligned} \frac{dy}{dt} &= \dot{y}, \\ \frac{d\dot{y}}{dt} &= -\zeta\dot{y} + \varphi(y, \dot{y}). \end{aligned}\right\} \qquad (10.19)$$

Конечно, при этом характер P- и N-совокупностей зависит от величины ζ. Но так как уравнения (10.19) не содержат y в явном виде, то эти совокупности дуг состоят из конгруэнтных кривых, сдвинутых вдоль оси y. На

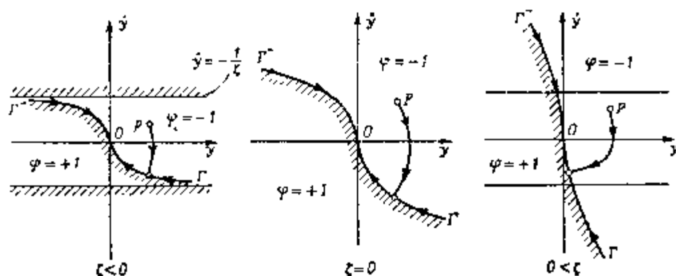
фиг. 95 изображено для каждого из трех видов значений ζ по одной типичной P-дуге и N-дуге, проходящих через начало координат. Случай $\zeta < 0$ отличается от двух остальных в том отношении, что система достигает начала координат только при условии, что начальное значение $\dot{y}$ заключено в границах между $-1/\zeta$ и $+1/\zeta$. Поэтому при $\zeta < 0$ задача о нахождении оптимальной функции включения имеет смысл только в том случае, когда начальная скорость $\dot{y}$ лежит внутри указанной полосы.



Фиг. 95

Обозначим через Γ ту часть P-дуги, проходящей через начало координат, которая лежит под осью y, а через Γ^- — отображение этой дуги относительно начала. Поэтому Γ^- представляет собой ту часть N-дуги, проходящей через начало, которая лежит над осью y. Дуги Γ и Γ^- образуют в совокупности некоторую кривую C. Бушау показал, что кривая C представляет собой оптимальную кривую переключения в том смысле, что над кривой C оптимальная функция $\varphi(y, \dot{y})$ переключения должна принимать значение -1, а под кривой C функция $\varphi(y, \dot{y})$ должна принимать значение $+1$. Графическая иллюстрация этого положения показана на фиг. 96. Физически процесс переключения реле осуществляется следующим образом: в любой точке p, расположенной над кривой C, вынуждающая функция должна равняться -1, и система движется вдоль N-дуги по направлению к кривой C переключения. На

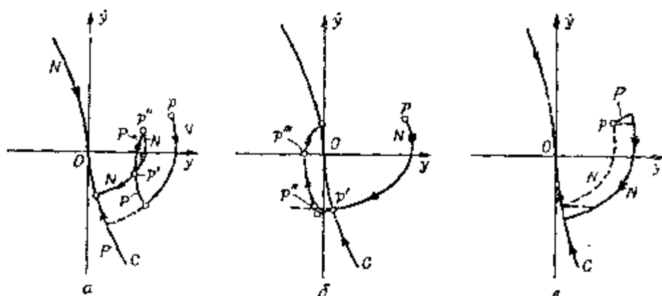
этой кривой происходит изменение значения вынуждающей функции, и последняя становится равной $+1$, а система движется вдоль кривой C к началу координат. Если начальная точка p расположена под кривой C, то вынуждающая функция должна равняться $+1$, и система движется вдоль P-дуги до пересечения с кривой C переключения. На этой кривой вынуждающая функция меняет свое значение на -1, и система движется к началу координат вдоль кривой C.



Фиг. 96

В том, что решение Бушау для оптимальной линии переключения является правильным, можно убедиться следующим образом. Прежде всего заметим, что участок кривой C служит последним участком траектории системы, в которой она достигает начала координат, так как кривая C является единственной из рассматриваемых траекторий, проходящих через начало. Предположим теперь, что начальная точка расположена над кривой C и что оптимальная траектория, определенная по методу Бушау, имеет вид, изображенный на фиг. 97, a, и состоит из одной N-дуги, идущей от точки p до кривой C, и следующей за ней дуги кривой C, идущей от точки пересечения с этой N-дугой до начала координат. Тогда, если переключение произведено слишком рано, то на части траектории, предшествующей встрече ее с кривой C, образуется одна угловая точка типа NP. Но, чтобы эта траектория достигла кривой C, необходимо выполнить еще одно переключение и образовать угловую точку типа PN. Если это переключение выполнено в точке p', где ско-

рость $\dot{y}$ еще сохраняет отрицательное значение, то появляется угловая точка типа PN под осью y, что приводит к нарушению правила построения канонической траектории. Продолжительность перемещения системы вдоль измененной таким образом траектории заведомо превышает продолжительность перемещения ее вдоль канонической траектории. Если же угловая точка типа PN совпадает



Фиг. 97

с p'', где скорость $\dot{y}$ положительна, то промежуток времени, необходимый для достижения положения равновесия, будет еще больше, так как в этом случае измененная траектория содержит замкнутый контур. Таким образом, преждевременное переключение является невыгодным. Построение, соответствующее запаздывающему переключению, приведено на фиг. 97,*б*. Так как траектории $p'0$ и $p''p'''$ эквивалентны [1]) и, таким образом, требуют одного и того же времени, то нетрудно заключить, что запаздывающее переключение также приводит к ухудшению процесса в системе. На фиг. 97,*в* приведена еще одна возможность, характеризующаяся тем, что первая часть траектории представляет собой P-дугу вместо N-дуги. Но из построения, выполненного на этой

---

[1]) Точка p'' имеет ту же ординату, что и p', а точка p''' ту же ординату, что и 0. Таким образом, отрезок $p''p'''$ равен отрезку $p'0$ и может быть получен перемещением отрезка $p'0$ влево. Следовательно, и времена перемещения по этим отрезкам одинаковы. — *Прим. ред.*

фигуре, следует, что и это видоизменение траектории системы приводит к ухудшению процесса в системе по сравнению с процессом, соответствующим оптимальной траектории. Эти соображения указывают на правильность выбора канонической траектории в качестве оптимальной кривой переключения.

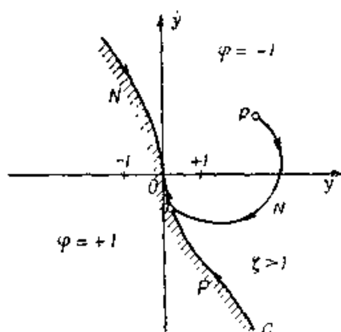
10.8. Оптимальная кривая переключения в линейных системах второго порядка. Бушау построил оптимальную кривую переключения для линейных систем второго порядка при $g(y, \dot{y}) = 2\zeta\dot{y} + y$ и любых вещественных значениях ζ. Мы ограничимся лишь изложением его результата, не приводя доказательства; однако в свете нашего рассуждения, относящегося к простому случаю, разобранному в предыдущем пункте, легко представить себе общий характер этого результата. P- и N-совокупности для данного вида функции $g(y, \dot{y})$ сводятся просто к семейству кривых, изображенных на фиг. 82—86, на которых начало координат следует сместить в точки $(+1, 0)$ и $(-1, 0)$ соответственно.

К нашему простому примеру ближе всего подходит система, в которой $\zeta > 1$. В этом случае кривая C переключения состоит из P-дуги, идущей от бесконечно удаленной точки до начала координат фазовой плоскости, и N-дуги, также идущей от бесконечно удаленной точки до начала. Как и раньше, над кривой C функция φ переключения принимает значение -1, а под этой кривой — значение $+1$. Оптимальная траектория, идущая от некоторой точки, находящейся на кривой C, до начала координат, имеет, таким образом, вид, изображенный на фиг. 98.

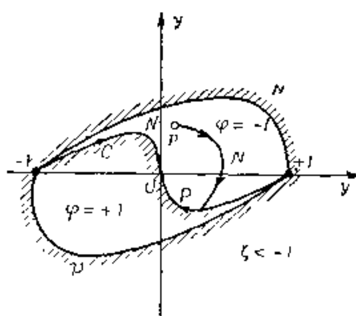
Так же, как и в нашем простом примере, при $\zeta < -1$ система может достичь начала координат только из точек, находящихся внутри определенной ограниченной области на фазовой плоскости, поскольку в отсутствии возмущающей функции система оказывается неустойчивой. Бушау указал, что граница этой области образована последовательным соединением P-дуги от точки $(+1,0)$ до точки $(-1,0)$ и N-дуги от точки $(-1,0)$ до точки $(+1,0)$ (фиг. 99). Задача об оптимальном законе переключения имеет смысл только в тех случаях, когда начальные точки лежат в этой области.

Оптимальная кривая C переключения состоит из P-дуги и N-дуги, причем движение системы на каждой из этих дуг происходит в направлении к началу координат. Над кривой C функция φ переключения равна —1, под этой кривой она равна +1. Образец оптимальной траектории, начинающейся в точке p, лежащей над кривой, приведен на фиг. 99.

При ζ 0 P- и N-совокупности сводятся к семействам окружностей с центрами в точках (+1,0) и (—1,0) соответственно. Оптимальная кривая C переключения (фиг. 100) состоит из последовательности полуокружностей



Фиг. 98

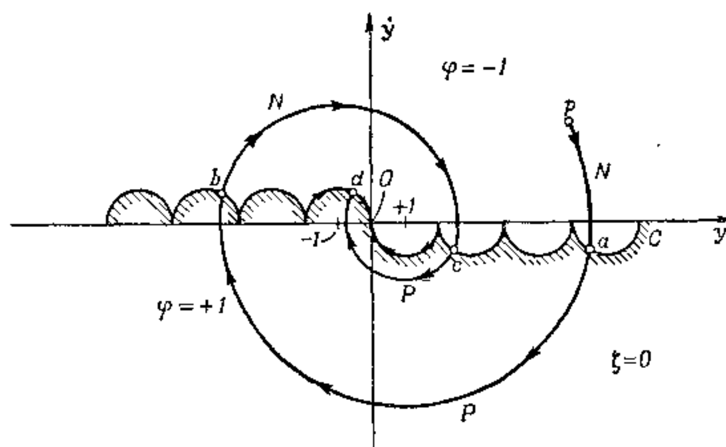


Фиг. 99

единичного радиуса, начинающейся в начале координат, причем эти полуокружности расположены вдоль оси y в обоих направлениях. Для положительных значений эти полуокружности расположены под осью y, а для отрицательных значений—над осью y. Над кривой C переключения функция φ переключения равна —1, а под ней φ равна +1. Таким образом, оптимальная траектория системы, начинающаяся в точке p на фиг. 100, следует вдоль N-дуги, представляющей собой дугу окружности с центром в точке (4,0). Когда эта траектория пересекается с кривой C в точке a, то она следует далее вдоль P-дуги, также являющейся дугой окружности, но с центром уже в точке (+1,0). После этого траектория системы вновь пересекает кривую C, на этот раз в точке b. Далее траектория вновь идет по очередной N-дуге вплоть до следующего пересечения с кривой C переключения и т. д. Последнее пересечение траектории с кривой C имеет место в точке d, и, начиная с этой

точки, траектория следует к началу координат вдоль кривой C. В рассматриваемом случае мы встречаемся со значительно более сложным законом переключения, чем при $\zeta > 1$.

Еще более сложным является случай движения изображающей точки системы по сходящимся спиралям, отвечающий неравенству $0 < \zeta < 1$. Как показал Бушау, в этом случае оптимальную кривую переключения следует строить следующим образом: проведем сперва из начала координат

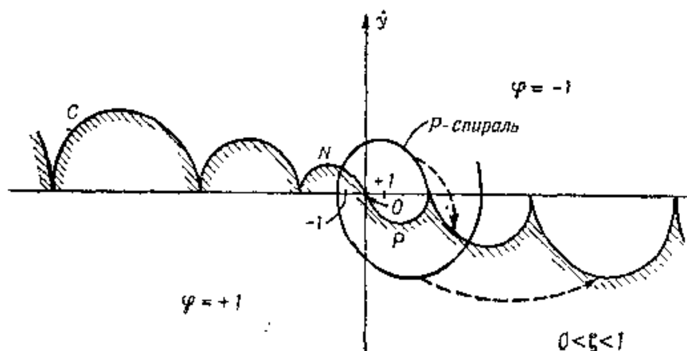


Фиг. 100

P-спираль в направлении, противоположном отсчету времени. Первой дугой кривой C будет начальная дуга этой P-спирали от начала до точки первого пересечения спирали с осью y. Затем построим отражения всех дуг, расположенных над осью y, относительно правой точки их пересечения с осью y. Все получившиеся в результате отражения дуги, расположенные под осью y, объединим в одну непрерывную кривую, передвигая дуги параллельно оси y до тех пор, пока они попарно не соединятся концами; эта непрерывная кривая исходит из начала координат и простирается вправо. Она и представляет собой половину оптимальной кривой переключения, соответствующую положительным значениям y.

Другая половина этой кривой, соответствующая отрицательным значениям y, получается затем в результате отражения относительно начала координат. Опять над кривой C переключения (фиг. 101) функция φ переключения равна —1, а под кривой φ равна +1. Общая картина в сильной степени похожа на картину, соответствующую случаю $\zeta=0$ и изображенную на фиг. 100; единственное различие состоит в замене дуг окружностей дугами спирали.

Остается разобрать последний случай, определяемый движением изображающей точки системы по расходящимся



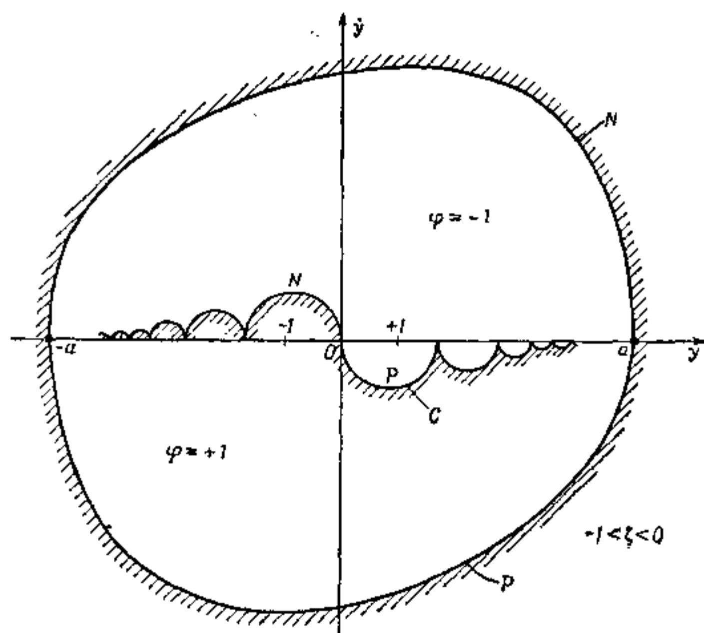
Фиг. 101

спиралям и отвечающий неравенству $-1<\zeta<0$. Метод построения оптимальной кривой переключения остается в точности таким же, как и в предыдущем случае, с той разницей, что теперь размеры последовательных дуг спиралей будут убывать, а не возрастать. Таким образом, длина кривой переключения является в этом случае конечной величиной, и эта кривая простирается вдоль оси от точки $(-a,0)$ до точки $(+a,0)$ (фиг. 102)[1].

Но так и должно быть: действительно, в данном случае система обладает отрицательным демпфированием, и, как и в случае, соответствующем построению на фиг. 99,

[1]) С математической точки зрения это утверждение следует обосновать, так как из убывания общего члена ряда, состоящего из положительных членов, еще не следует сходимость этого ряда. Здесь это утверждение основывается лишь на физической аналогии. — *Прим. перев.*

фазовая траектория системы может достичь начала координат лишь при том условии, если ее начальная точка заключена внутри некоторой области около начала. В действительности граница этой области состоит из P-дуги на участке от точки $(a,0)$ до точки $(-a,0)$ и N-дуги на участке от точки $(-a,0)$



Фиг. 102

до точки $(a,0)$. Оптимальная функция φ переключения равна -1 в точках фазовой плоскости, расположенных над кривой C, и равна $+1$ в точках, расположенных под этой кривой.

Замкнутые кривые, ограничивающие на фиг. 96 и 99 области возможного оптимального закона переключения, представляют собой, очевидно, предельные циклы, причем переключения происходят в точках $\dot{y}=0$. Каждый из этих предельных циклов характеризует периодическое движение рассматриваемой релейной следящей системы. Однако в такой же мере очевиден и неустойчивый характер периодических движений: при малейшем возмущении фазовые

траектории этих систем будут или спиралеобразно наматываться на начало координат или расходиться. Поэтому в реальных системах эти периодические колебания не могут иметь места[1]).

Во всех рассмотренных случаях отчетливо выступает следующее свойство найденного нами решения задачи об оптимальном законе переключения реле: оптимальная функция φ переключения принимает значение -1 в первом квадранте фазовой плоскости и значение $+1$ в третьем квадранте. Переписав систему уравнений (10.19) в виде одного уравнения второго порядка

$$\frac{d^2y}{dt^2} = -2\zeta\frac{dy}{dt} - y + \varphi\left(y, \frac{dy}{dt}\right),$$

мы сможем очень легко уяснить себе сущность этого свойства решения нашей задачи. Назначение релейного устройства в системе состоит в том, чтобы возвращать систему к состоянию равновесия $y=0$, и при этом в возможно более короткое время. Когда обе функции y и dy/dt положительны, этой цели можно достичь, придавая второй производной d^2y/dt^2 [характеризующей кривизну кривой $y(t)$] отрицательные значения, как можно большие по абсолютной величине, т. е. функция φ должна принимать значение -1. Когда обе функции y и dy/dt отрицательны, вторая производная d^2y/dt^2 должна быть положительной и как можно большой, т. е. функция φ должна принимать значение $+1$. Это интуитивное рассуждение совпадает с нашими выводами о строении оптимальной функции переключения. Когда же функции y и $\dot{y}$ имеют различные знаки, то найти оптимальное значение функции φ уже не так просто, так как в этом случае быстрота убывания координаты y в функции времени t зависит от сложного взаимного влияния функций y и $\dot{y}$. Поэтому вклад Бушау в теорию релейных систем состоит в установлении оптимальных значений φ именно в таких областях изменения y и $\dot{y}$, т. е. во втором и четвертом квадрантах фазовой плоскости. Но из этих соображений

---

[1]) Такие предельные циклы называются неустойчивыми.—*Прим. ред.*

становится очевидным, что оптимальная кривая C переключения должна лежать во втором и четвертом квадрантах.

В системах более высокого порядка и системах более чем с одной степенью свободы изображение состояния системы на фазовой плоскости становится невозможным. В этих условиях приходится переходить к фазовому пространству многих измерений. По аналогии с задачей, рассмотренной выше, можно ожидать, что решение задачи об оптимальном законе переключения приведется к разысканию оптимальных поверхностей переключения в фазовом пространстве. Однако эта задача еще не решена[1]); известна лишь единственная попытка в этом направлении, принадлежащая Кангу и Фетту[2]).

10.9. Системы с несколькими законами управления. Что будет происходить с системой, когда она в результате целенаправленных переключений релейного звена приблизится к состоянию равновесия, характеризуемому началом координат на фазовой плоскости?

Ясно, что если вынуждающая сила будет продолжать действовать на систему с той же интенсивностью, с которой она на нее действовала в моменты, непосредственно предшествовавшие приходу в начало координат, то система будет снова удаляться от желаемого состояния равновесия. Но, как только она начнет удаляться от начала координат, переключающее (релейное) устройство, которым снабжена система, будет действовать так, чтобы принуждать систему возвращаться обратно к началу координат. В результате система будет быстро приближаться к началу координат из любого возмущенного состояния, после чего в ней установятся высокочастотные колебания («рысканье») около этого равновесного состояния.

Если в системе допустимы малые отклонения от состояния покоя (начала координат), то от вибрационных движений около начала можно избавиться путем отключения вынуждающей функции каждый раз, когда система нахо-

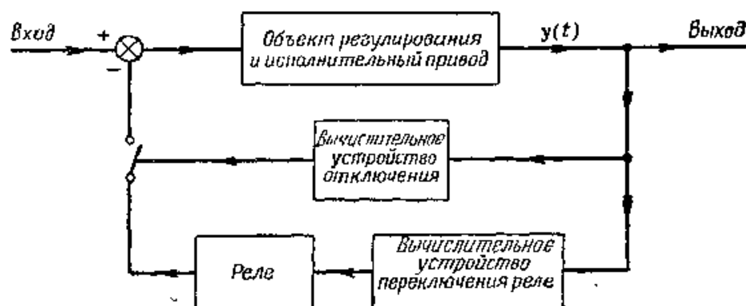
---

[1]) Эта задача решена для широкого класса систем в n-мерном пространстве. См. [49, 50], а также [16].—*Прим. ред.*

[2]) Kang C. L., Fett G. H., Journ. Appl. Phys, **24**, 38—41 (1953).

дится вблизи состояния равновесия, т. е. когда y и $\dot{y}$ пренебрежимо малы. Таким образом, в системе действуют два закона управления: при больших отклонениях системы ею управляет закон переключения через посредство выходного воздействия релейного звена; при малых отклонениях такое воздействие отключено. К необходимости подобных воздействий с различными законами управления в следящих системах мы вновь вернемся в последующих главах.

Поскольку оптимальная кривая переключения нелинейна, оптимальный закон переключения нельзя физи-



Фиг. 103

чески осуществить посредством простой линейной цепи. В самом деле, значение выходной величины $y(t)$, измеренное с помощью соответствующего звена, необходимо «переработать» в нелинейном звене, т. е. в вычислительном устройстве, выполненном таким образом, чтобы оно вырабатывало сигнал переключения, поступающий на вход релейного звена, в соответствии с оптимальной кривой переключения. Далее, в системе должно иметься другое вычислительное устройство, предназначенное вырабатывать сигнал, от которого производится отключение выхода релейного звена от системы, как только величины y и $\dot{y}$ становятся достаточно малыми. В результате блок-схема такой релейной следящей системы имеет вид, показанный на фиг. 103[1]).

---

[1]) О нелинейных обратных связях см. также статью [55] и сборник [44], стр. 296—319.—*Прим. перев.*

В более сложных системах, которые мы рассмотрим в дальнейших главах, включение в систему вычислительных устройств в общем случае представляется необходимым. Однако по идее это устройство не вносит ничего нового; в самом деле, ведь корректирующий контур, применяемый в обычных следящих системах с целью видоизменения передаточной функции, также представляет собой вычислительное устройство.

Но в этих более простых системах вычислительные операции можно выполнять с помощью линейной цепи, например RC-цепи. К более подробному исследованию вычислительных устройств мы вернемся в гл. XIII.

Глава XI

НЕЛИНЕЙНЫЕ СИСТЕМЫ

Система называется нелинейной, если ее выход не представляет собой линейной функции от ее входа. Простым примером нелинейной системы служит релейная следящая система. В гл. VI мы описали общий метод физической линеаризации любой нелинейной системы автоматического управления или слежения; любую нелинейную систему можно заставить действовать, как линейную, если превратить ее в систему колебательного управления. В предыдущей главе мы изложили метод исследования следящих систем, содержащих нелинейное звено, отработка которого не чувствительна к частоте входа. Эти методы расчета нелинейных следящих систем применимы к весьма разнообразным нелинейным задачам, возникающим в технической практике, и достаточно мощны для решения задач синтеза обычных систем.

С другой стороны, как отмечено в последних разделах гл. X, задача об оптимальном использовании нелинейной характеристики системы для повышения ее качества является в общем случае гораздо более трудной по сравнению с задачей обеспечения устойчивости. По существу, в этом направлении пока что заложено лишь скромное начало. Поэтому в настоящее время не представляется возможным дать удовлетворительное исследование в области нелинейных следящих систем общего типа. Более того, и саму задачу следовало бы, вероятно, сформулировать более прямым образом и в ином плане. Именно, вместо исследования свойств некоторой данной системы следовало бы задать наперед требуемые свойства системы и затем разыскать необходимую для этого нелинейную характеристику.

Такое направление мы рассмотрим в гл. XIV. Цель же настоящей главы весьма ограничена: мы укажем только на

некоторые возможности целенаправленного использования характеристик нелинейной системы.

11.1. Релейные следящие системы с нелинейной обратной связью. Ограничиваясь рассмотрением случая малых отклонений системы от состояния равновесия, можно сильно упростить результаты Бушау об оптимальной кривой переключения в релейной следящей системе, изложенные в п. 10.8. Обращаясь к материалу этого пункта, мы видим, что вблизи начала координат кривая C переключения приближенно описывается уравнением

$$\dot{y}|\dot{y}| = -2y. \tag{11.1}$$

Это означает, что с помощью нелинейного закона переключения, определяемого соотношением (11.1), можно улучшить качество системы по сравнению с системой, в которой применяется линейный закон переключения (п. 10.6). Обозначив через x вход системы, а через y — ее выход, перепишем соотношения (11.1) в виде

$$a^2\dot{y}|\dot{y}| = x - y \tag{11.2}$$

или

$$\operatorname{sign}(x-y)\sqrt{|x-y|} = a\dot{y}, \tag{11.3}$$

где a — постоянная [1]).

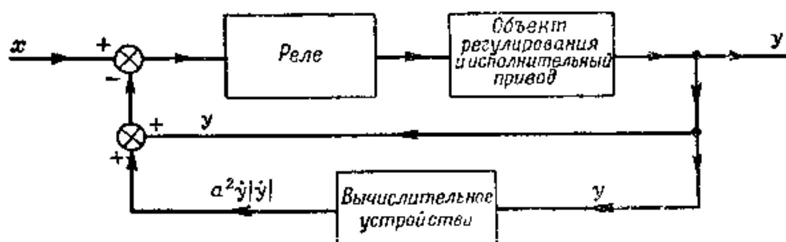
Блок-схема релейной следящей системы, работающей по закону переключения (11.2), представлена на фиг. 104. Обратная связь, действующая по закону $\dot{y}|\dot{y}|$, была предложена Аттли и Хэммондом [2]) в целях улучшения качества простой релейной следящей системы [3]). Вычислительное устройство здесь необходимо только для выработки сигнала $a^2\dot{y}|\dot{y}|$ и поэтому может быть относительно про-

[1]) Здесь $x-y$ есть погрешность системы, а $\dot{y}$ — ее производная (с точностью до знака), если считать $x = \text{const}$. — *Прим. ред.*

[2]) Аттли А. М. и Хэммонд П. Х., статья в сборнике «Автоматическое регулирование», М., 1954, стр. 296.

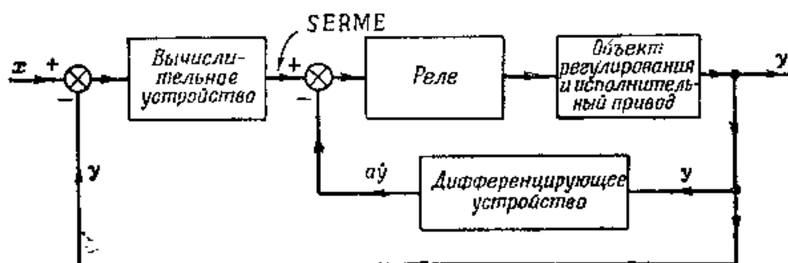
[3]) Значительно ранее такая обратная связь была предложена Д. И. Марьяновским и Д. В. Свечарником. См. по этому поводу [46]. — *Прим. ред.*

стым. Равносильный закон переключения, определяемый соотношением (11.3), также можно осуществить с помощью обратной связи по $\dot{y}$ и по $\operatorname{sign}(x-y)\sqrt{|x-y|}$. Так как разность $x-y$ равна рассогласованию в системе, систему



Фиг. 104

можно назвать системой с управлением по квадратному корню из модуля рассогласования (система SERME). Блок-схема системы изображена на фиг. 105. И в этом случае вычислительное устройство является сравнительно простым. Эту систему предложил Д. С. Вест[1]).



Фиг. 105

Хотя эти нелинейные связи в релейных следящих системах сравнительно просты, они не поддаются строгому аналитическому исследованию. В самом деле, наши доводы в их пользу являются только допустимыми, но не исчерпывающими. Окончательную же доработку всякой конкретной системы, в которой применяются эти принципы, следует производить экспериментальным путем.

[1]) Сборник «Автоматическое регулирование», М., 1954, стр. 315.— *Прим. перев.*

11.2. Системы с малой нелинейностью. Если система, обладающая n степенями свободы, является нелинейной, то вместо линейного дифференциального уравнения типа (2.3) мы будем иметь уравнение

$$a_n \frac{d^n y}{dt^n} + a_{n-1} \frac{d^{n-1} y}{dt^{n-1}} + \ldots + a_1 \frac{dy}{dt} + a_0 y + \mu f\left(y, \frac{dy}{dt}, \ldots, \frac{d^{n-1} y}{dt^{n-1}}\right) = x(t), \quad (11.4)$$

где $a_n, \ldots, a_0$ и μ — постоянные, $x(t)$ — вход, $y(t)$ — выход, а f — нелинейная функция своих переменных. Сумма первых $n+1$ членов в левой части уравнения (11.4) представляет собой левую часть дифференциального уравнения (2.3). Все нелинейные зависимости системы выражаются через посредство последнего слагаемого в левой части уравнения (11.4). Говоря о малой нелинейности, мы имеем в виду, что постоянная μ мала по сравнению с коэффициентами $a_0, \ldots, a_n$ [1]).

В случае системы с малой нелинейностью мы можем попытаться построить решение в виде формального разложения в ряд по степеням μ:

$$y(t) = y^{(0)}(t) + \mu y^{(1)}(t) + \mu^2 y^{(2)}(t) + \ldots \quad (11.5)$$

Подставив разложение (11.5) в (11.4) и приравняв коэффициенты при одинаковых степенях μ, получим

$$a_n \frac{d^n y^{(0)}}{dt^n} + a_{n-1} \frac{d^{n-1} y^{(0)}}{dt^{n-1}} + \ldots + a_1 \frac{dy^{(0)}}{dt} + a_0 y^{(0)} = x(t), \quad (11.6)$$

$$a_n \frac{d^n y^{(1)}}{dt^n} + a_{n-1} \frac{d^{n-1} y^{(1)}}{dt^{n-1}} + \ldots + a_1 \frac{dy^{(1)}}{dt} + a_0 y^{(1)} = -f\left(y^{(0)}, \ldots, \frac{d^{n-1} y^{(0)}}{dt^{n-1}}\right) \quad (11.7)$$

и другие уравнения для коэффициентов при высших членах разложения. Приближение нулевого порядка опреде-

---

[1]) Постоянную μ нельзя сравнивать со всеми коэффициентами $a_0, \ldots, a_n$ одновременно, так как размерности этих коэффициентов не могут совпадать, какова бы ни была размерность переменной y. Такое сравнение можно провести лишь при переходе к безразмерному времени $\tau = t/t^*$, где t^* — постоянная, имеющая размерность времени, например $t^* = 1$ сек. — *Прим. перев.*

ляется, таким образом, линейным уравнением (2.3). Но для нас более важно то обстоятельство, что первая поправка, обусловленная нелинейностью, определяется дифференциальным уравнением (11.7), характеристическое уравнение для которого совпадает с характеристическим уравнением для дифференциального уравнения нулевого приближения. Другими словами, если из линейного приближения обнаруживается, что система демпфирована и обладает другими желаемыми качествами следящей системы, то и поправка $y^{(1)}(t)$ первого порядка будет также обладать этими свойствами. Больше того, поскольку малый параметр μ умножается на $y^{(1)}(t)$ во втором слагаемом правой части равенства (11.5), то и поправки на нелинейность малы. Следовательно, влияние малой нелинейности на систему, обладающую удовлетворительным «качеством», не приведет к существенному изменению процессов в системе по сравнению с процессами в линеаризованной системе. Поэтому, коль скоро задача рассматривается в плане технического приближения, мы можем обращаться с такими системами, как с линейными. В этом и заключается причина столь широкого и успешного распространения линейной теории следящих систем, несмотря на присутствие малых нелинейностей даже во всех «линейных» системах[1]).

С другой стороны, если линейная система очень слабо демпфирована, то, как мы знаем, в системе может возникнуть резонанс. Иными словами, даже если порядок величины входа $x(t)$ равен единице, то выход $y^{(0)}(t)$ системы линейного приближения, описываемой уравнением (11.6), может иметь порядок гораздо более высокий, чем единица. При этих обстоятельствах величина

$$\mu f\left(y^{(0)}, \ldots, \frac{d^{n-1}y^{(0)}}{dt^{n-1}}\right),$$

характеризующая влияние нелинейностей, может быть того же порядка, как и некоторые из линейных членов даже при малом μ. Иначе говоря, здесь наше фор-

---

[1]) Однако встречаются случаи, когда даже малые нелинейности создают дополнительные существенные эффекты в системах (люфт, гистерезис, иногда зона нечувствительности и т. д.). — *Прим. ред.*

мальное разложение искомого решения в ряд, введенное на стр. 257, не является более справедливым, и следует ожидать сильных эффектов даже от малой нелинейности, если система обладает очень слабым демпфированием.

В последующих разделах мы дадим краткое описание поистине калейдоскопического поведения[1]) нелинейных систем. Подробное изучение этих явлений нелинейной механики можно найти в прекрасных книгах Минорского[2]) и Стокера[3]).

11.3. Явление скачка. Как уже отмечалось в п. 10.2, самовозбуждающиеся колебания системы называются «мягкими», если они начинаются автоматически уже при очень малых отклонениях системы от состояния равновесия; если же эти колебания возникают лишь при больших отклонениях от состояния равновесия, они называются «жесткими».

В некоторых случаях коэффициенты дифференциального уравнения могут зависеть от некоторого параметра λ системы. Если при некотором определенном значении $\lambda=\lambda_0$, которое мы назовем критическим, состояние равновесия, до тех пор устойчивое, становится неустойчивым, то при этом появляется предельный цикл, т. е. возникают автоколебания. Это явление графически изображено на фиг. 106, *а* для случая мягкого самовозбуждения и на фиг. 106, *б* для случая жесткого самовозбуждения. В первом случае, при возрастании параметра λ, в системе не будет происходить никаких изменений до тех пор, пока λ не достигнет величины λ_0, при которой характер состояния равновесия меняется от устойчивого к неустойчивому с одновременным появлением устойчивого предельного цикла, причем амплитуда соответствующих автоколебаний с увеличением λ растет. При убывании λ явление протекает в точности в обратном направлении, и при $\lambda=\lambda_0$ предельный цикл исчезает. В системах с жестким самовозбуждением имеет место иная картина (фиг. 106, *б*). Автоколебания возникают внезапно при $\lambda=\lambda_0$, имея конечную

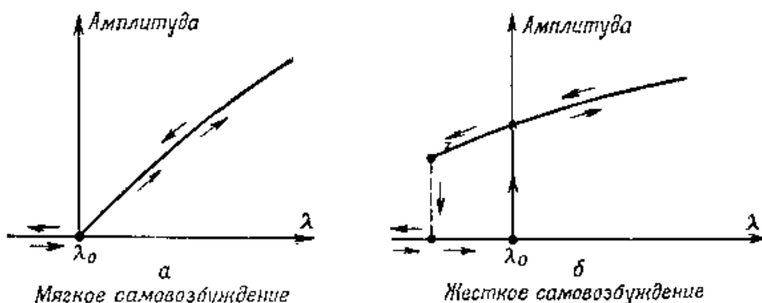
---

[1]) В смысле крайнего разнообразия свойств. — *Прим. ред.*

[2]) Minorsky N., Introduction to nonlinear mechanics, Ann Arbor, Mich., 1947.

[3]) Стокер Дж., Нелинейные колебания в механических и электрических системах, М., 1952.

амплитуду; с дальнейшим ростом λ амплитуда этих колебаний увеличивается. Если затем начать уменьшать параметр λ, то при возвращении его к значению $\lambda=\lambda_0$ автоколебания и, следовательно, предельный цикл не исчезают при этом значении λ; они исчезают позже, при дальнейшем перемещении вдоль кривой, в точке, в которой



Фиг. 106

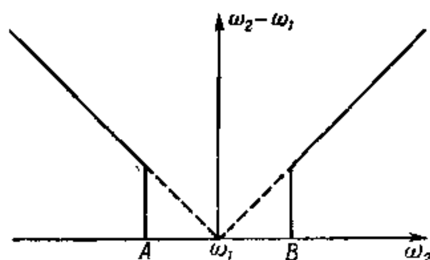
амплитуда вновь мгновенно меняется скачком от конечного значения до нуля. Это явление скачка связано с гистерезисом в поведении системы.

11.4. Деление частот. Если к нелинейной системе приложено периодическое входное воздействие, состоящее из двух колебаний с частотами ω_1 и ω_2, то выход системы будет содержать не только эти частоты и их гармоники, но также дополнительный спектр так называемых комбинированных тонов $m\omega_1 \pm n\omega_2$, где m и n суть целые числа. Так, например, если на вход нелинейного проводника подать напряжение, изменяющееся по закону $x = x_0(\cos\omega_1 t + \cos\omega_2 t)$, и если в этом проводнике протекает ток, связанный с напряжением x зависимостью $y = a_1x + a_2x^2 + a_3x^3$, то выход $y(t)$ проводника будет содержать колебания со следующими частотами: ω_1, ω_2, $2\omega_1$, $2\omega_2$, $3\omega_1$, $3\omega_2$, $\omega_1+\omega_2$, $\omega_1-\omega_2$, $2\omega_1+\omega_2$, $2\omega_1-\omega_2$, $\omega_1+2\omega_2$ и $\omega_1-2\omega_2$. Первые шесть частот определяют обычные гармоники, вторые шесть представляют собой комбинированные тоны, обусловленные нелинейностью проводника. Некоторые из них выше исходных частот ω_1 и ω_2, дру-

гие—ниже. Эти более низкие частоты называются *субгармониками*, а процесс, в результате которого они возникают, называется *делением* частот.

Очевидно, что если частоты ω_1 и ω_2 очень близки одна к другой, то разность $\omega_1 - \omega_2$ гораздо меньше любой из исходных частот. Если к тому же систему можно стабилизировать на этой низкой частоте, то можно получить субгармоники порядка 0,01 от входной частоты и даже еще меньшего. Еще более низкие частоты можно получить последовательным соединением нескольких таких систем таким образом, чтобы субгармонический выход одной системы служил входом следующей системы.

11.5. Явление захватывания. Пусть в нелинейной системе произошло самовозбуждение и установились автоколебания с частотой ω_1; тогда при подаче на вход системы колебаний с частотой ω_2, слегка отличной от частоты ω_1, мы могли бы ожидать появления на выходе



Фиг. 107

обоих колебаний с частотами ω_1 и ω_2 одновременно и, кроме того, биений с частотой $\omega_2 - \omega_1$, возникающих в результате нелинейных взаимовлияний в системе. В действительности, процесс будет протекать согласно графику, изображенному на фиг. 107. Колебания с частотой $\omega_2 - \omega_1$ внезапно исчезают, как только частота ω_2 достигает так называемой *зоны синхронизации AB*. В этом интервале существуют только колебания с частотой ω_2, и процессы в системе протекают так, как если бы колебания с исходной частотой ω_1 захватывались колебаниями с изменяющейся частотой ω_2.

Это явление *захватывания частот* впервые объяснил ван-дер Поль; затем его идеи развили другие исследователи. Пусть рассматриваемая система принадлежит к системам второго порядка и описывается дифференциальным уравнением

$$\frac{d^2y}{dt^2} - \alpha\frac{dy}{dt} + \gamma\left(\frac{dy}{dt}\right)^3 + \omega_1^2 y = B\omega_1^2 \sin\omega_2 t, \qquad (11.8)$$

где α, γ и B — положительные постоянные. При $B=0$ система обладает отрицательным демпфированием при колебаниях с малыми амплитудами и положительным демпфированием при колебаниях с большими амплитудами. Таким образом, существует амплитуда, при которой в системе могут установиться автоколебания. Более того, если постоянные α и γ обе малы, в системе должны возникнуть самовозбуждающиеся колебания с частотой, близкой к ω_1. Ван-дер Поль показал, что, коль скоро ω_2 близко к ω_1, решение уравнения (11.8) можно представить в форме

$$y(t) = b_1(t)\sin\omega_2 t + b_2(t)\cos\omega_2 t, \qquad (11.9)$$

где $b_1(t)$ и $b_2(t)$ — медленно изменяющиеся функции t, такие, например, что [1])

$$\left|\frac{db_1}{dt}\right| \ll |\omega_2 b_1(t)| .$$

Подставляя решение (11.9) в уравнение (11.8) и удерживая лишь члены не выше первой степени, мы можем написать уравнения, определяющие функции b_1 и b_2, в виде

$$\left.\begin{aligned} \frac{db_1}{dt} &= f_1(b_1, b_2; \omega_2), \\ \frac{db_2}{dt} &= f_2(b_1, b_2; \omega_2). \end{aligned}\right\} \qquad (11.10)$$

Мы получили систему автономных уравнений первого порядка. Поэтому их можно решить по методу изоклин таким же образом, как решается и уравнение второго порядка с помощью построений на фазовой плоскости. Как показывает более подробное исследование, в некотором диапазоне значений ω_2, вблизи ω_1, система уравнений (11.10) имеет устойчивый узел на плоскости координат

[1]) См. приложение в конце книги [1]. — *Прим. перев.*

b_1, b_2. Таким образом, независимо от начальных значений переменных b_1 и b_2 система стремится к определенным установившимся значениям этих переменных, соответствующим такому узлу. Следовательно, колебания с частотой $\omega_2 - \omega_1$ и не могут возникнуть. При величинах ω_2, выходящих за границы этого диапазона, на плоскости координат b_1, b_2 существует устойчивый предельный цикл, обусловливающий явление, изображенное на фиг. 107.

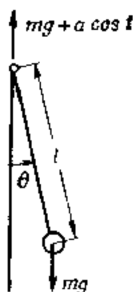
11.6. Асинхронное возбуждение и подавление колебаний. В некоторых нелинейных системах существует возможность возбуждать или прекращать некоторые колебания с частотой ω с помощью других колебаний совершенно иной частоты ω_1.

В первом случае это явление называется *асинхронным возбуждением*, во втором - *асинхронным подавлением*. Для уяснения существа данного явления вспомним, что один только факт существования предельного цикла в фазовом пространстве рассматриваемой системы еще не означает безусловного наступления автоколебаний. Чтобы автоколебания фактически имели место, предельный цикл должен быть устойчивым в том смысле, что система будет стремиться возвратиться к этому предельному циклу, даже если она подвергнется возмущению и переместится в фазовом пространстве в точку, удаленную от этого цикла. Неустойчивый же предельный цикл в реальной системе осуществить невозможно. В свете этих соображений легко заметить, что при некоторых обстоятельствах возникновение колебаний нового вида может привести к созданию или разрушению условий устойчивости колебаний другого вида.

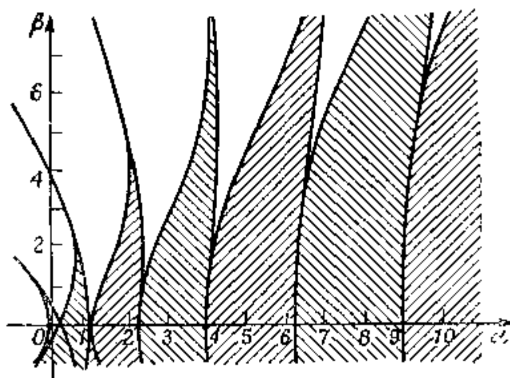
В первом случае мы встречаемся с явлением асинхронного возбуждения колебаний, во втором – с явлением асинхронного подавления колебаний. Слово «асинхронный» употребляется лишь для того, чтобы подчеркнуть, что частоты ω и ω_1 не связаны каким-либо определенным соотношением [1]).

[1]) Имеется в виду, что эти частоты не связаны соотношением с целочисленными коэффициентами, не зависящим от параметров системы (например, ω_1 в общем случае не есть кратное ω_2 и т. д.). — *Прим. перев.*

11.7. Параметрическое возбуждение и демпфирование. Уже с давних пор было известно, что если в колебательной системе периодически изменять какой-либо параметр с некоторой частотой ω, то система начнет колебаться с частотой $\omega/2$. Рэлей продемонстрировал это явление на примере камертона, к одной из пластин которого был прикреплен конец длинной проволоки. Если пластина камертона колеблется с частотой ω, то возникают боковые колебания проволоки с частотой $\omega/2$. Подобная же



Фиг. 108



Фиг. 109

картина имеет место в случае колебаний математического маятника, находящегося под действием синусоидальной силы, приложенной к верхнему концу стержня (фиг. 108). Обозначив через θ малое угловое перемещение маятника от вертикали и предположив, без ограничения общности, что частота синусоидальной силы равна единице, напишем дифференциальное уравнение движения в виде[1])

$$ml\frac{d^2\theta}{dt^2} + (mg + a\cos t)\,\theta = 0,$$

где m — масса колеблющейся точки, g — ускорение силы тяжести, l — длина маятника, а a — амплитуда периоди-

[1]) Задача о движении маятника при колебаниях точки подвеса рассмотрена в работе [63]. — *Прим. перев.*

ческой силы. Это уравнение можно представить в форме

$$\frac{d^2\theta}{dt^2}+(\alpha+\beta\cos t)\,\theta=0. \tag{11.11}$$

Здесь α равно g/l, а β равно a/ml. Уравнение (11.11) справедливо также для случая обращенного маятника, в котором масса находится над точкой подвеса, если считать g отрицательным. Поэтому для обычного маятника величина α положительна, а для обращенного — отрицательна. Уравнение (11.11) представляет собой линейное уравнение, но с периодически изменяющимся коэффициентом, характеризующим силу, приложенную к точке подвеса маятника. Таким образом, эту систему можно рассматривать как систему с периодическим параметром.

Уравнение (11.11) представляет собой хорошо известное дифференциальное уравнение Матье[1]). Устойчивость решения зависит от значений постоянных α и β. Действительно, на плоскости параметров α и β можно выделить область устойчивости и область неустойчивости, как это показано на фиг. 109 (заштрихована область устойчивости). Таким образом, в случае обычного маятника ($\alpha>0$) система устойчива в отсутствии периодической силы (при $\beta=0$), что, конечно, хорошо известно. Однако интересно отметить, что при надлежащем выборе β систему можно превратить в неустойчивую. Тогда маятник будет раскачиваться с возрастающей амплитудой до тех пор, пока влияние нелинейных явлений не приведет к установлению постоянной амплитуды некоторой довольно значительной величины. В этом и заключается явление *параметрического возбуждения*. При отрицательных значениях α, т. е. в случае обращенного маятника без наличия периодической силы, приложенной к точке опоры ($\beta=0$), система, разумеется, неустойчива. Но при значениях β, принадлежащих к некоторому узкому диапазону, систему можно стабилизировать. Это явление называется *параметрическим демпфированием*.

Параметрическое возбуждение или параметрическое демпфирование могут иметь место в любой системе с периодически изменяющимися параметрами. Это явление,

---

[1]) См. [64]. — *Прим. перев.*

а также любое из рассмотренных выше нелинейных явлений можно использовать в целях достижения желаемого качества системы автоматического регулирования. Многие из этих явлений уже широко применяются в большом количестве различных звеньев, из которых состоят системы автоматического регулирования. Однако эти нелинейные звенья преимущественно играют роль вспомогательных приспособлений, и их применяют скорее на основании опыта и данных эксперимента, чем на основании теоретического исследования. Вопрос о построении нелинейной характеристики всей системы автоматического регулирования в совокупности все еще остается неисследованным. Материал, изложенный нами в предыдущих разделах, служит лишь указанием на большие возможности в этой области[1]).

---

[1]) О синтезе нелинейных оптимальных систем см. [78].— *Прим. ред.*

Глава XII

ЛИНЕЙНЫЕ СИСТЕМЫ С ПЕРЕМЕННЫМИ КОЭФФИЦИЕНТАМИ

Единственной системой с переменными коэффициентами, которую мы подробно рассматривали до сих пор, был маятник, находящийся под действием периодической силы, приложенной в точке подвеса; его колебаниями мы занимались в связи с явлением параметрического возбуждения и демпфирования. Во всех же других случаях дифференциальные уравнения, описывавшие рассмотренные нами системы, не содержали коэффициентов, явно зависящих от времени. Однако мы показали в гл. I, что поведение линейных систем, описываемых уравнениями с переменными коэффициентами, может коренным образом отличаться от поведения систем, описываемых уравнениями с постоянными коэффициентами. В настоящей главе мы вновь обратимся к этому вопросу и довольно подробно рассмотрим типичную, но простую систему: баллистическую ракету ближнего действия. Мы покажем на этом примере, что задачу об устойчивости системы с переменными коэффициентами, вообще говоря, нельзя решать теми же методами, какими решают задачу об устойчивости линейных систем с постоянными коэффициентами. Здесь мало отметить, что метод преобразования Лапласа и понятие передаточной функции бесполезны для нашей задачи[1]); более того, мы сталкиваемся с необходимостью изменить существо самого подхода к решению такой задачи.

Мы начнем с изучения движения баллистической ракеты, стабилизируемой с помощью хвостового оперения, в период работы ее двигателя, создающего тягу. При этом мы обра-

---

[1]) Однако имеются попытки распространения преобразования Лапласа на линейные системы с переменными параметрами [65]. — *Прим. перев.*

тим особое внимание на угловые отклонения оси ракеты от требуемого угла стрельбы, возникающие под действием возмущений во время запуска ракеты, и на последующее демпфирующее действие ее оперения. Общую задачу динамики баллистической ракеты подробно изучали многие исследователи в различных странах во время второй мировой войны. Итоги американских работ в этой области изложены в книге Россера, Ньютона и Гросса[1]). Работа, проделанная в Англии, изложена в книге Ренкина[2]). Французские исследования на ту же тему отражены в статье Каррьера[3]). Мы разберем здесь эту задачу в весьма упрощенной постановке, имея целью лишь оттенить существенно интересные моменты в плане изучения линейных систем с переменными коэффициентами.

12. 1. Движение баллистической ракеты на активном участке траектории. В случае полета баллистической ракеты, снабженной стабилизирующим хвостовым оперением, взаимодействие между составляющими движениями в вертикальной и горизонтальной плоскостях пренебрежимо мало, т. е. движение ракеты в горизонтальной плоскости приводит к появлению лишь пренебрежимо малых аэродинамических сил, действующих в вертикальной плоскости. Поэтому мы не упустим никаких характерных свойств ракеты, если ограничимся рассмотрением явлений, протекающих в вертикальной плоскости, и изучением движения ракеты только в этой плоскости. Поскольку мы разбираем полет ракеты ближнего действия, мы будем считать поверхность земли плоской. Пусть v обозначает абсолютную величину скорости ракеты[4]), θ—угол наклона ее вектора скорости к горизонту, φ—угол наклона оси ракеты к горизонту[5]) (фиг. 110). При этом угол атаки α выражается с помощью соотношения

$$\alpha=\varphi-\theta. \tag{12.1}$$

---

[1]) Россер Дж., Ньютон Р. и Гросс Г., Математическая теория полета неуправляемых ракет, М., 1950.

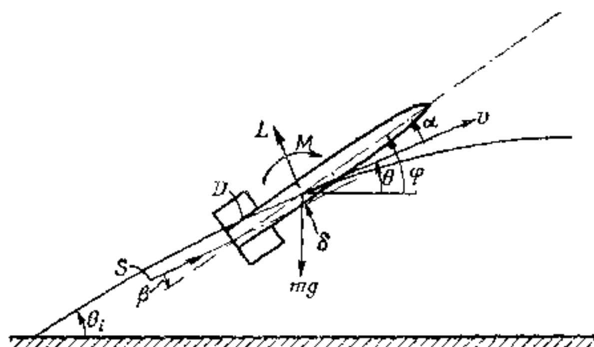
[2]) Ренкин Р. А., Математическая теория движения неуправляемых ракет, М., 1951.

[3]) Carrière P., Mem. artillerie franç., 25, 253—360 (1951).

[4]) Здесь v—аэродинамическая скорость.—*Прим. перев.*

[5]) То есть угол тангажа.—*Прим. перев.*

Пусть m — масса ракеты и g — постоянное ускорение силы тяжести. Сила тяжести равна mg и направлена вертикально вниз. Введем далее обозначения аэродинамических сил: L — подъемная сила, D — лобовое сопротивление, где сила L направлена перпендикулярно к направлению движения, а сила D параллельно направлению движения. Момент аэродинамических сил обозначим через M. Аэродинамические силы приложены к центру масс ракеты[1]).



Фиг. 110

Тяга S двигателя ракеты, в данном случае постоянная, может действовать под некоторым углом β к оси ракеты (β — угол установки двигателя) и иметь плечо δ относительно ее центра масс[2]) (фиг. 110). Тогда уравнения движения в проекциях на касательную и на нормаль к траектории имеют вид

$$m\frac{dv}{dt} = S\cos(\alpha - \beta) - mg\sin\theta - D, \tag{12.2}$$

$$mv\frac{d\theta}{dt} = S\sin(\alpha - \beta) - mg\cos\theta + L. \tag{12.3}$$

---

[1]) Точка приложения аэродинамической силы $L+D$ в плоском потоке определяется теоремой С. А. Чаплыгина [66, 67]. Аэродинамическую силу можно, конечно, считать приложенной к центру масс с соответствующим изменением момента M по правилу механики о перемене центра приведения системы сил. К тому же центр масс перемещается при горении относительно ракеты, и поэтому момент M будет зависеть также и от t [68]. — *Прим. перев.*

[2]) В общем случае во время горения $\delta = \delta(t)$. — *Прим. перев.*

Если k есть радиус инерции ракеты относительно боковой оси[1]), проходящей через центр масс ракеты, то уравнение моментов имеет вид

$$mk^2 \frac{d^2\varphi}{dt^2} = S\delta - M. \tag{12.4}$$

В уравнении (12.4) мы пренебрегли так называемым демпфирующим моментом реактивной струи[2]), поскольку его величина мала по сравнению с величиной восстанавливающего момента, создаваемого хвостовым оперением.

Аэродинамические сила и момент зависят от угла α атаки. Но при неточной установке хвостового оперения относительно корпуса ракеты подъемная сила и аэродинамический момент не обращаются в нули при угле атаки, равном нулю. Чтобы учесть это явление, введем так называемый угол γ нулевой подъемной силы, так что L и M равны нулю не при $\alpha = 0$, а при $\alpha = \gamma$. Пусть ρ — плотность воздуха, d — диаметр корпуса ракеты; тогда коэффициент K_L подъемной силы, коэффициент K_D лобового сопротивления и коэффициент K_M продольного аэродинамического момента получат следующие выражения:

$$L = K_L \rho v^2 d^2 \sin(\alpha - \gamma), \tag{12.5}$$

$$D = K_D \rho v^2 d^2, \tag{12.6}$$

$$M = K_M \rho v^2 d^3 \sin(\alpha - \gamma). \tag{12.7}$$

При стрельбе баллистическими ракетами ближнего действия вершина траектории такой ракеты расположена на небольшой высоте над поверхностью земли; поэтому плотность ρ воздуха можно принять постоянной. Кроме того, наибольшая скорость ракеты достаточно мала для того, чтобы коэффициенты K_L, K_D и K_M можно было считать постоянными, т. е. не зависящими от изменения числа Маха вдоль траектории. Больше того, для ракет ближнего действия масса горючего составляет лишь небольшую часть массы ракеты, и, следовательно, массу m ракеты

---

[1]) В общем случае $k = k(t)$ как вследствие выгорания массы, так и вследствие перемещения центра масс в теле ракеты. — *Прим. перев.*

[2]) См. [69]. — *Прим. перев.*

можно без серьезной погрешности считать постоянной. Таким образом, траектория ракеты определяется уравнениями (12.1) — (12.7) с учетом введенных упрощающих предположений.

Время горения в ракетах этого типа очень мало, порядка 1/5 сек. Отсюда следует, что ускорение S/m должно быть непременно очень большим. В реальных устройствах тяга S настолько велика, что по сравнению с ней силы тяготения и лобового сопротивления пренебрежимо малы. К тому же угловое отклонение направления тяги от направления полета, т. е. угол $\alpha - \beta$, никогда не бывает велико. Поэтому уравнение (12.2) в нулевом приближении приводится к виду

$$\frac{dv}{dt} = \frac{S}{m}. \tag{12.8}$$

Таким образом, нулевым приближением траектории служит прямая линия, наклоненная к горизонту под углом, равным начальному углу θ_i (фиг. 110). Движение вдоль этой прямой совершается с постоянным ускорением S/m. Следовательно, если через z обозначить путь, проходимый ракетой вдоль этой прямой, то закон движения определится уравнением

$$v^2 = \frac{2S}{m} z. \tag{12.9}$$

Если запуск ракеты происходит без начальной скорости, то z и представляет собой истинное расстояние от точки запуска. Если же ракета выпущена с некоторой начальной скоростью, то z представляет собой уже не расстояние ракеты от точки запуска, а ее расстояние от некоторой другой точки, расположенной впереди точки запуска. В силу равенства (12.9), будем иметь

$$\frac{d}{dt} = \frac{dz}{dt}\frac{d}{dz} = v\frac{d}{dz} = \sqrt{\frac{2S}{m}z}\,\frac{d}{dz}; \tag{12.10}$$

Исходя из нулевого приближения решения, основанного на равенстве (12.10), мы можем подсчитать первое приближение траектории ракеты. К этим вычислениям мы теперь и переходим.

12.2. Линеаризованные уравнения траектории. Так как во время горения отклонение истинной траектории от ее нулевого приближения никогда не бывает большим, мы можем в уравнениях (12.3) и (12.4) всюду заменить скорость v и ее производную по времени их приближениями, определенными с помощью соотношений (12.9) и (12.10). Кроме того, в силу малости угла $\alpha - \beta$, $\sin(\alpha - \beta)$ можно заменить дугой $\alpha - \beta$; подобным же образом вместо $\cos\theta$ можно подставить $\cos\theta_i$. Мы пренебрежем также подъемной силой L, ибо она мала по сравнению с нормальной составляющей тяги и с весом ракеты. С учетом указанных упрощений уравнения (12.3) и (12.4) примут вид

$$2z\frac{d\theta}{dz} = (\alpha - \beta) - \frac{mg}{S}\cos\theta_i \tag{12.11}$$

и

$$2z\frac{d^2\varphi}{dz^2} + \frac{d\varphi}{dz} = \frac{\theta}{k^2} - \frac{8\pi^2}{\sigma^2}z(\alpha - \gamma), \tag{12.12}$$

где σ — постоянная, определяемая равенством

$$\sigma^2 = 4\pi^2\frac{k^2 m}{K_M \rho d^3}. \tag{12.13}$$

Величину σ, имеющую, очевидно, размерность длины, можно принять за характеристическую длину возмущенного движения ракеты и рассматривать ее как длину волны возмущенной траектории. Следовательно, уравнения (12.1), (12.11) и (12.12) суть линеаризованные уравнения относительно трех неизвестных α, θ и φ. Линеаризация проведена в предположении, что отклонение траектории от прямолинейной, соответствующей идеальному значению θ_i угла бросания, является малой величиной.

Мы можем исключить θ и φ из этих уравнений, чтобы получить единственное уравнение относительно α. Для этого разделим обе части уравнения (12.11) на $2\sqrt{z}$ и продифференцируем уравнение, полученное в результате этого деления, по переменной z. Мы получим

$$\sqrt{z}\frac{d^2\theta}{dz^2} + \frac{1}{2\sqrt{z}}\frac{d\theta}{dz} = \frac{1}{2\sqrt{z}}\frac{d\alpha}{dz} - \frac{1}{4z\sqrt{z}}\left(\alpha - \beta - \frac{gm}{S}\cos\theta_i\right).$$

Разделим теперь обе части уравнения (12.12) на $2\sqrt{z}$ и вычтем из результата только что полученное уравнение. Тогда, используя уравнение (12.1), придем к уравнению

$$\sqrt{z}\frac{d^2\alpha}{dz^2}+\frac{1}{\sqrt{z}}\frac{d\alpha}{dz}+\left(\frac{4\pi^2\sqrt{z}}{\sigma^2}-\frac{1}{4z\sqrt{z}}\right)\alpha=$$
$$=\frac{\delta}{2k^2\sqrt{z}}+\frac{4\pi^2\sqrt{z}}{\sigma^2}\gamma-\frac{1}{4z\sqrt{z}}\left(\beta+\frac{gm}{S}\cos\theta_i\right).$$

Обращаясь к последнему уравнению, мы обнаружим, что это дифференциальное уравнение, с помощью которого определяется устойчивость баллистической ракеты, не принадлежит к уравнениям с постоянными коэффициентами. В самом деле, это уравнение можно привести к обычной форме уравнения Бесселя путем введения безразмерного расстояния ξ, определяемого соотношением

$$\xi=\frac{2\pi z}{\sigma}, \tag{12.14}$$

где σ — «длина волны», определенная посредством соотношения (12.13). Рассматриваемое уравнение получит тогда вид

$$\frac{d^2\alpha}{d\xi^2}+\frac{1}{\xi}\frac{d\alpha}{d\xi}+\left[1-\frac{(1/2)^2}{\xi^2}\right]\alpha=$$
$$=\gamma+\left(\frac{\delta\sigma}{4\pi k^2}\right)\frac{1}{\xi}-\frac{1}{4}\left(\beta+\frac{gm}{S}\cos\theta_i\right)\frac{1}{\xi^2}. \tag{12.15}$$

После того как угол α будет найден, θ можно определить путем интегрирования следующего уравнения, полученного с помощью уравнения (12.11):

$$\frac{d\theta}{d\xi}=\frac{1}{2\xi}\left(\alpha-\beta-\frac{mg}{S}\cos\theta_i\right). \tag{12.16}$$

Независимая переменная z или ξ представляет собой не время, а расстояние. Однако поскольку на основании равенства (12.9) переменные z и, следовательно, ξ представляют собой монотонные функции от t, мы можем заменить независимую переменную t независимой переменной ξ, не изменяя при этом устойчивости системы в том смысле, что если система устойчива по ξ, то она будет устойчива и по t; это означает, что в обоих случаях отклонение

траектории от идеальной прямолинейной траектории будет убывать с ростом ξ или t. Поэтому задачу об исследовании устойчивости системы можно с полным правом решать, пользуясь уравнением, в котором независимой переменной служит величина ξ, возрастающая с течением времени от ее начального значения ξ_0. Начальное значение ξ_0 равно нулю, если запуск производится без начальной скорости.

12.3. Устойчивость баллистической ракеты. Для выяснения вопроса об устойчивости ракеты мы должны решить уравнения (12.15) и (12.16) при определенных начальных условиях и затем определить, стремится ли к нулю угол атаки α (или, точнее говоря, угловое отклонение $\theta - \theta_i$ возмущенной траектории от невозмущенной) по мере возрастания ξ. Уравнение (12.15) является уравнением Бесселя порядка 1/2. Решение соответствующего однородного уравнения составляется, таким образом, из функций Бесселя порядков 1/2 и —1/2. Эти функции можно, однако, выразить через элементарные функции. В самом деле, перепишем уравнение (12.15) в виде

$$\frac{d^2\zeta}{d\xi^2} + \zeta = Q(\xi), \tag{12.17}$$

где

$$\zeta = \sqrt{\xi}\,\alpha \tag{12.18}$$

и

$$Q(\xi) = \gamma\sqrt{\xi} + \left(\frac{\delta\sigma}{4\pi k^2}\right)\frac{1}{\sqrt{\xi}} - \frac{1}{4}\left(\beta + \frac{gm}{S}\cos\theta_i\right)\frac{1}{\xi^{3/2}}. \tag{12.19}$$

Поэтому частные решения однородного уравнения для ζ сводятся просто к $\sin\xi$ и $\cos\xi$.

Начальные условия, т. е. условия в момент схода ракеты с направляющих устройств, выражаются равенствами

$$\left.\begin{aligned} v &= v_0, \\ \theta &= \theta_0, \\ \alpha &= \alpha_0 \\ \text{и}\quad \frac{d\varphi}{dt} &= \left(\frac{d\varphi}{dt}\right)_0. \end{aligned}\right\} \tag{12.20}$$

Индекс 0 указывает на то, что значение соответствующей переменной берется в момент времени $t=0$. На основании начальных условий задачи подсчитаем начальные значения переменных ξ и ζ:

$$\xi_0 = \frac{2\pi}{\sigma}\frac{m}{2S}v_0^2 = \frac{\pi m v_0^2}{\sigma S} \tag{12.21}$$

и

$$\zeta_0 = \sqrt{\xi_0}\,\alpha_0. \tag{12.22}$$

С помощью уравнения (12.16) найдем, что при $t=0$

$$\sqrt{\xi_0}\left(\frac{d\theta}{d\xi}\right)_0 = \frac{\alpha_0 - \beta - (mg/S)\cos\theta_i}{2\sqrt{\xi_0}}.$$

Но $\theta = \varphi - \alpha$ и, следовательно,

$$\sqrt{\xi_0}\left(\frac{d\alpha}{d\xi}\right)_0 + \frac{1}{2\sqrt{\xi_0}}\alpha_0 = \left(\frac{d\zeta}{d\xi}\right)_0 =$$

$$= \sqrt{\xi_0}\left(\frac{d\varphi}{d\xi}\right)_0 + \frac{\beta + (gm/S)\cos\theta_i}{2\sqrt{\xi_0}},$$

или

$$\left(\frac{d\zeta}{d\xi}\right)_0 = \sqrt{\xi_0}\,\frac{\sigma\,(d\varphi/dt)_0}{2\pi v_0} + \frac{\beta + (gm/S)\cos\theta_i}{2\sqrt{\xi_0}}. \tag{12.23}$$

С помощью преобразованных таким образом начальных условий мы можем выписать решение уравнения (12.17); выразив ζ через α по формуле (12.18), получим

$$\alpha(\xi, \xi_0) = \frac{1}{\sqrt{\xi}}\cos(\xi - \xi_0)\left[\zeta_0 - \int_{\xi_0}^{\xi}\sin(\eta - \xi_0)\,Q(\eta)\,d\eta\right] +$$

$$+ \frac{1}{\sqrt{\xi}}\sin(\xi - \xi_0)\left[\left(\frac{d\zeta}{d\xi}\right)_0 + \int_{\xi_0}^{\xi}\cos(\eta - \xi_0)\,Q(\eta)\,d\eta\right], \tag{12.24}$$

где Q — вынуждающая функция, определенная соотношением (12.19). Так как функция $Q(\eta)$ содержит половинные степени от η, то интегралы, входящие в решение (12.24), являются интегралами Френеля. После того как угол α определен, переменная θ находится с помощью квадрату-

ры на основании уравнения (12.16):

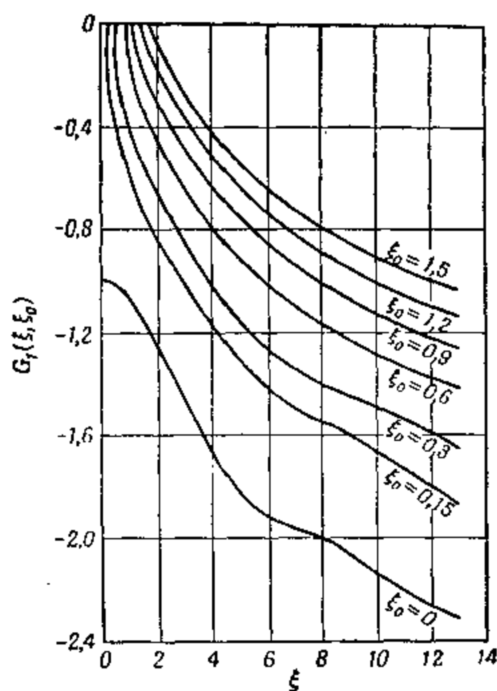
$$\theta-\theta_i=(\theta_0-\theta_i)-\frac{1}{2}\left(\beta+\frac{mg}{S}\cos\theta_i\right)\ln\frac{\xi}{\xi_0}+$$
$$+\frac{1}{2}\int\limits_{\xi_0}^{\xi}\frac{\alpha(\eta,\ \xi_0)}{\eta}\,d\eta. \qquad (12.25)$$

В этом выражении для θ мы можем провести дальнейшее разделение членов, выражающих возмущения различной природы, имеющие место при разгоне ракеты на направляющих. Для этого перепишем выражение (12.2) в виде суммы членов, каждый из которых отражает определенное возмущение. Таким образом,

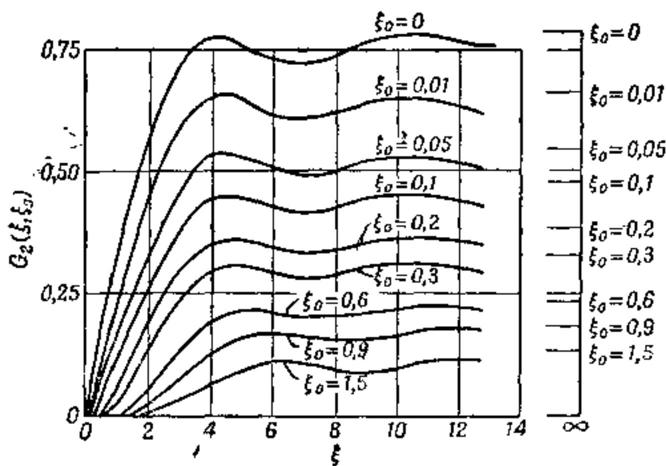
$$\theta-\theta_i=(\theta_0-\theta_i)+\left(\beta+\frac{mg}{S}\cos\theta_i\right)G_1(\xi,\ \xi_0)+$$
$$+\frac{\delta\sigma}{2\pi k^2}G_2(\xi,\ \xi_0)-\gamma\left[G_1(\xi,\ \xi_0)+G_3(\xi,\ \xi_0)\right]+$$
$$+\alpha_0 G_3(\xi,\ \xi_0)+\frac{1}{2}\frac{\sigma}{\pi v_0}\left(\frac{d\varphi}{dt}\right)_0 G_4(\xi,\ \xi_0). \qquad (12.26)$$

Первое слагаемое в правой части равенства (12.26) характеризует влияние на полет начального отклонения угла наклона траектории. Второе слагаемое отражает влияние угла β и поля земного тяготения. Третье слагаемое выражает воздействие момента тяги относительно центра масс. Четвертое слагаемое определяет составляющую отклонения траектории, обусловленную влиянием угла установки хвостового оперения. Влияние на полет начального угла атаки сказывается через посредство пятого слагаемого. Последнее слагаемое учитывает влияние начальной угловой скорости ракеты. Каждая из функций G_1, G_2, G_3 и G_4, зависящих от переменной ξ и параметра ξ_0, представляет собой линейную форму от интегралов Френеля. Россер и его сотрудники назвали эти функции «функциями расчета ракеты» и дали в своей книге таблицы этих функций. Фиг. 111—114, на которых приведены графики функций G_1, G_2, G_3 и G_4, также заимствованы из этой книги.

Рассматривая графики на фиг. 111—114, мы увидим, что при достаточно больших значениях ξ каждая из функций G_1, G_2, G_3 и G_4 приблизительно сохраняет постоянное

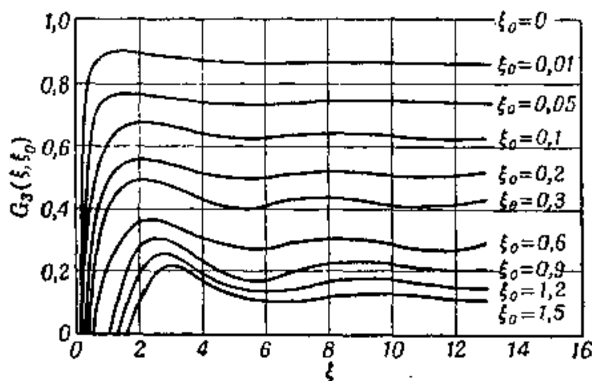


Фиг. 111

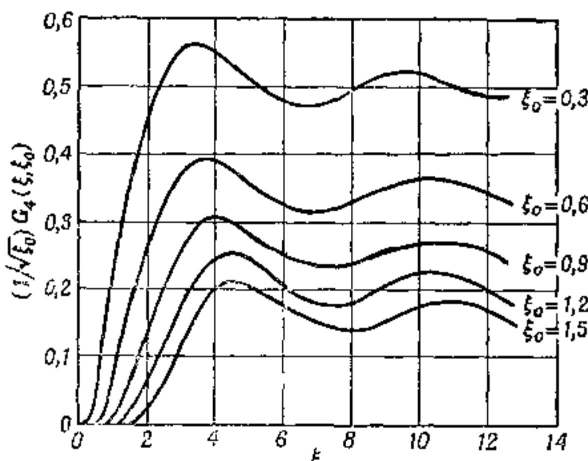


Фиг. 112

значение. Таким образом, возмущения в полете полностью не подавляются. Первое и два последних слагаемых в правой



Фиг. 113



Фиг. 114

части равенства (12.26) характеризуют влияние начальных возмущений при сходе ракеты с направляющих. При больших значениях ξ эти слагаемые принимают конечные ненулевые значения. Другие слагаемые в правой части равен-

ства (12.26) представляют собой «выход от входного воздействия», или вынуждающие функции. Эти слагаемые также принимают некоторые конечные значения при больших ξ. Хуже ведет себя функция G_1: при больших значениях ξ она приближенно равна $\ln \xi$ и, следовательно, возрастает неограниченно при возрастании ξ. Поэтому, если мы будем оценивать полет ракеты с точки зрения установленного ранее критерия устойчивости системы, т. е. будем считать систему устойчивой, если начальные возмущения затухают, а выход системы при «разумных» ограничениях величины входной функции остается ограниченным, то полет баллистической ракеты является неустойчивым. С другой стороны, функции, определяющие общее решение однородного уравнения, соответствующего основному уравнению (12.15), суть функции Бесселя, которые исчезают при больших значениях переменной и таким образом как бы указывают на присущую системе устойчивость. Если мы попытаемся подойти к оценке систем с переменными коэффициентами с точки зрения опыта, приобретенного при исследовании систем, описываемых линейными дифференциальными уравнениями с постоянными коэффициентами, то поведение системы с переменными коэффициентами окажется для нас явно непонятным. Но это обстоятельство лишь указывает на неприменимость к системам с переменными коэффициентами понятий, относящихся к системам с постоянными коэффициентами. Здесь требуется новый подход к задаче об устойчивости и автоматическом регулировании. Этот вопрос мы рассмотрим в следующем пункте.

12. 4. Устойчивость и автоматическое регулирование систем с переменными коэффициентами. Как показывают наши предыдущие рассуждения, для случая линейной системы с постоянными коэффициентами можно гарантировать удовлетворительное качество системы в отношении ее устойчивости и других критериев регулирования, если общее решение однородного уравнения (без вынуждающей функции или входа) достаточно быстро затухает. Поэтому, несмотря на то, что вынуждающая функция или вход могут меняться от задачи к задаче в значительной степени, критерий качества системы и, следовательно, ее расчет основываются на изучении решения однородного уравнения, т. е. на изучении

собственных колебаний системы. В этом заключается основной принцип теории следящих систем. С этой точки зрения метод исследования систем с помощью передаточных функций, основанный на преобразовании Лапласа, представляет собой лишь технический прием. В принципе, например, классический метод решения однородных уравнений ровно настолько же хорош, насколько хорош и метод годографа корней Ивэнса.

Исследование, проведенное выше в этой главе, со всей определенностью показывает, что в линейных системах с переменными коэффициентами затухание всех собственных колебаний в системе еще не обеспечивает удовлетворительного качества этой же системы при воздействии на нее вынуждающей силы. При некоторых входах (или вынуждающих функциях) выход может даже стать неограниченным, несмотря на затухающий характер собственных колебаний системы. Таким образом, качество системы нельзя оценить без знания входной функции. В свете этого требования конкретного задания входной функции, а также трудности фактического нахождения решения неоднородного дифференциального уравнения с переменными коэффициентами задача о построении общей теории устойчивости и регулирования в таких системах может показаться безнадежной. Однако мы должны проводить различие между вычислительными трудностями и реальными трудностями создания общей теории. Трудности вычислительного порядка можно преодолеть с помощью быстродействующих электрических вычислительных устройств[1]), и потому их не следует принимать за реальные трудности. Уяснив себе это обстоятельство, мы придем к выводу о том, что требование задания конкретной входной функции при исследовании системы сводится, по существу, к требованию того, чтобы было указано назначение системы; прежде чем приступить к разработке системы, мы должны знать, что мы хотим от системы и при каких обстоятельствах. Коль скоро принят именно этот путь, общая задача об устойчивости системы уже отпадает, так как если система рассчитана на выполнение определенного удовлетворяющего нас

---

[1]) В настоящее время наибольшее распространение получили электронные вычислительные устройства.—*Прим. ред.*

действия, то это свойство ее само по себе является достаточным. Для линейных систем с переменными параметрами общую теорию регулирования можно построить именно на этой основе, применяя классическую баллистическую теорию возмущений, что мы и рассмотрим в следующей главе. Обозревая весь предыдущий материал, мы можем сказать, что теория обычных следящих систем является теорией, посвященной общим методам проектирования систем частного вида. Теория возмущений, которую мы изложим в следующей главе, представляет собой теорию, посвященную частным задачам проектирования систем более общего вида.

Глава XIII

РАСЧЕТ СИСТЕМ АВТОМАТИЧЕСКОГО РЕГУЛИРОВАНИЯ НА ОСНОВЕ ТЕОРИИ ВОЗМУЩЕНИЙ

Назначение баллистической теории возмущений состоит в подсчете величин, определяющих поведение баллистического снаряда вблизи так называемой нормальной траектории[1]). Нормальной траекторией называется определенная траектория, отвечающая известным начальным условиям, заданному закону тяги, данному состоянию атмосферы и заданной программе изменения подъемной силы и лобового сопротивления. Если же истинные условия полета несколько отличаются от этих заданных условий или же если летящий объект претерпел отклонение от его нормальной траектории из-за случайного порыва ветра, то траектория объекта будет отличаться от нормальной. Но если эти возмущающие влияния малы, то возмущенная траектория будет проходить вблизи от нормальной траектории и различие между возмущенной и нормальной траекториями будет оставаться незначительным. Именно эта близость возмущенной траектории к теоретической, известной траектории, служит основанием для линеаризации дифференциальных уравнений движения по возмущенной траектории. Возмущенное движение описывается при этом линейной системой уравнений с переменными коэффициентами; эти коэффициенты изменяются во времени в силу того, что с течением времени меняются условия полета объекта.

Первоначальное назначение баллистической теории возмущений состояло в том, чтобы подсчитывать малые изменения траектории снаряда, вызываемые отклонениями веса оболочки снаряда от его стандартного значения, изменениями атмосферных условий, действием ветра и т. д. Однако с появлением современных больших и быстродействующих вычислительных машин наметилась тенденция вычислять

---

[1]) Эта траектория называется также невозмущенной.—*Прим. перев.*

все близкие возмущенные траектории порознь и благодаря этому отпала надобность в баллистической теории возмущений. Но задача расчета регулирующих устройств для линейных систем с переменными коэффициентами как раз и является той задачей, к которой можно применить баллистическую теорию возмущений, что мы проиллюстрируем в настоящей главе на примере задачи об управлении ракетой дальнего действия. Такая задача была рассмотрена Дреником[1]). Однако мы разберем здесь эту задачу в более полном объеме, включая вопрос об автоматическом управлении полетом этой ракеты[2]).

13.1. Уравнения движения ракеты. Чтобы не усложнять вычислений, предположим, что ракета движется в экваториальной плоскости вращающейся Земли (фиг. 115). Плоское

Фиг. 115

движение возможно благодаря отсутствию в экваториальной плоскости кориолисовой силы. Полет ракеты рассмат-

[1]) Дреник Р., Вопросы ракетной техники, сборник переводов, М., 5, 1952.

[2]) Tsien N. S., Adamson T. C., Knuth E. L., Journ. Am. Rocket Soc., 22, 192—199 (1952).

ривается в системе координат, связанной с вращающейся Землей, т. е. сама эта система вращается с угловой скоростью Ω вращения Земли. Положение летящей ракеты в экваториальной плоскости в любой момент времени t определяется полярным радиусом-вектором r и полярным углом θ, отсчитываемым от направления r в точке запуска ракеты. Пусть r_0 обозначает радиус Земли и g—постоянное значение ускорения силы тяжести на поверхности Земли без учета центробежной силы, обусловленной вращением Земли. Пусть, далее, R и Θ обозначают приходящиеся на единицу массы ракеты составляющие силы тяги и аэродинамических сил, приложенных к ракете, в радиальном и трансверсальном направлениях соответственно. При этих условиях уравнения движения центра масс ракеты имеют вид

$$\left.\begin{aligned} &\frac{dr}{dt}=\dot{r},\\ &\frac{d\theta}{dt}=\dot{\theta},\\ &\frac{d\dot{r}}{dt}=R+r(\dot{\theta}\pm\Omega)^2-g\left(\frac{r_0}{r}\right)^2,\\ &r\frac{d\dot{\theta}}{dt}=\Theta-2\dot{r}(\dot{\theta}\pm\Omega). \end{aligned}\right\} \qquad (13.1)$$

Знак «плюс» внутри скобок в правых частях двух последних уравнений отвечает полету в восточном направлении, а знак «минус» — в западном.

Силы R и Θ определяются тягой S, подъемной силой L и лобовым сопротивлением D. Пусть W — вес ракеты в данный момент времени, соответствующий ускорению g силы тяжести, а V — скорость ракеты относительно воздуха (воздушная скорость). Для дальнейшего удобно ввести параметры Σ, Λ и Δ, определяемые посредством соотношений

$$\Sigma=\frac{Sg}{W}; \quad \Lambda=\frac{Lg}{WV}; \quad \Delta=\frac{Dg}{WV}. \qquad (13.2)$$

Мы будем предполагать, что скорость w ветра направлена горизонтально и лежит в экваториальной плоскости, и будем считать ее положительной, если ветер встречный. Скорость

w рассматривается как функция высоты r. Обозначим через v_r радиальную скорость, а через v_θ — трансверсальную, т. е.

$$\left.\begin{aligned} v_r &= \dot{r}, \\ v_\theta &= r\dot{\theta}. \end{aligned}\right\} \qquad (13.3)$$

С помощью последних соотношений скорость V ракеты относительно воздуха определяется равенством

$$V^2 = \dot{r}^2 + (r\dot{\theta} + w)^2. \qquad (13.4)$$

Обозначив через β угол наклона линии действия силы тяги к горизонту, выразим радиальную и трансверсальную составляющие в виде

$$\left.\begin{aligned} R &= \Sigma \sin\beta + (v_\theta + w)\Lambda - v_r\Delta, \\ \Theta &= \Sigma \cos\beta - v_r\Lambda - (v_\theta + w)\Delta. \end{aligned}\right\} \qquad (13.5)$$

Обозначим через N результирующий момент всех приложенных сил относительно центра масс ракеты, отнесенный к ее моменту инерции, и напишем уравнение, определяющее угловое ускорение:

$$\frac{d\dot{\beta}}{dt} = \frac{d\dot{\theta}}{dt} + N. \qquad (13.6)$$

Для того чтобы иметь уравнения движения ракеты в окончательном виде, необходимо задать подъемную силу L, лобовое сопротивление D и результирующий момент m всех приложенных сил относительно центра масс в функции времени. Как принято в аэродинамике, выразим силы L и D через посредство коэффициента C_L подъемной силы и коэффициента C_D лобового сопротивления в следующей форме:

$$\left.\begin{aligned} L &= \frac{1}{2}\rho V^2 A C_L, \\ D &= \frac{1}{2}\rho V^2 A C_D, \end{aligned}\right\} \qquad (13.7)$$

где ρ — плотность воздуха, зависящая от r, а A — характерная площадь, например площадь крыла ракеты. В рас-

сматриваемой задаче, поскольку движение ракеты ограничено экваториальной плоскостью, положение ее на траектории, существенно важное для подсчета аэродинамических коэффициентов, определяется углом атаки, т. е. углом между линией действия силы тяги и вектором воздушной скорости ракеты (фиг. 115). Управление же полетом ракеты осуществляется путем изменения угла γ отклонения руля высоты. Таким образом, переменными, от которых зависят коэффициенты C_L и C_D, являются α и γ. Кроме того, аэродинамические коэффициенты зависят от числа Рейнольдса Re и числа Маха M. Обозначив через a скорость звука в воздухе, являющуюся функцией r, определим число Маха отношением

$$M = \frac{V}{a}. \tag{13.8}$$

Пусть l — характерный размер ракеты, а μ — коэффициент вязкости воздуха, также являющийся функцией r. Число Рейнольдса определяется соотношением

$$\mathrm{Re} = \frac{\rho V l}{\mu}. \tag{13.9}$$

Следовательно,

$$\begin{aligned} C_L &= C_L(\alpha,\ \gamma,\ M,\ \mathrm{Re}), \\ C_D &= C_D(\alpha,\ \gamma,\ M,\ \mathrm{Re}). \end{aligned} \tag{13.10}$$

Допустим, что линия действия силы тяги проходит через центр масс ракеты; таким образом, сила тяги не дает момента относительно центра масс. Поскольку следует ожидать, что угловое движение ракеты на активном участке траектории будет медленным, можно считать, что демпфирующий момент ракеты пренебрежимо мал. Единственный момент, приложенный к ракете, сводится при этих условиях к аэродинамическому моменту m. Этот аэродинамический момент можно выразить через посредство коэффициента C_M продольного момента [1]):

$$m = \frac{1}{2}\rho V^2 A l C_M. \tag{13.11}$$

---

[1]) Соотношения (13.7) и (13.11) отличаются от соотношений (12.5)—(12.7) лишь формой группировки множителей в их правых частях, а также тем, что в (12.5)—(12.7) не учитывается зависимость сил и момента от Re и M, а $\gamma = 0$. — *Прим. перев.*

Как и другие аэродинамические коэффициенты, коэффициент $C_{\rm M}$ продольного момента зависит от четырех переменных α, γ, M и Re:

$$C_{\rm M} = C_{\rm M}(\alpha,\ \gamma,\ M,\ \mathrm{Re}). \tag{13.12}$$

Обозначим через I значение момента инерции ракеты относительно ее боковой центральной оси в момент времени t. Тогда величина N, входящая в равенство (13.6), определится отношением

$$N = \frac{m}{I}. \tag{13.13}$$

Пользуясь введенными выше обозначениями, мы можем написать систему уравнений движения ракеты в виде

$$\frac{dr}{dt} = v_r, \qquad \frac{d\theta}{dt} = \frac{v_\theta}{r}, \qquad \frac{d\beta}{dt} = \dot\beta,$$

$$\frac{dv_r}{dt} = \Sigma \sin\beta + (v_\theta + w)\Lambda - v_r\Delta + r\left(\frac{v_\theta}{r} \pm \Omega\right)^2 - g\left(\frac{r_0}{r}\right)^2 = F,$$

$$\frac{dv_\theta}{dt} = \Sigma\cos\beta - v_r\Lambda - (v_\theta + w)\Delta - 2v_r\left(\frac{v_\theta}{r} \pm \Omega\right) + \frac{v_\theta v_r}{r} = G,$$

$$\frac{d\dot\beta}{dt} = \frac{1}{r}\left[\Sigma\cos\beta - v_r\Lambda - (v_\theta + w)\Delta - 2v_r\left(\frac{v_\theta}{r} \pm \Omega\right)\right] + N = H. \tag{13.14}$$

Уравнения (13.14) представляют собой систему уравнений первого порядка относительно шести неизвестных функций r, θ, β, v_r, v_θ и $\dot\beta$. Для определения этих функций необходимо задать шесть начальных значений, т. е. значений этих функций в момент $t = 0$. Кроме того, необходимо задать закон изменения во времени силы тяги S, веса W и момента I инерции. Для определения аэродинамических сил должен быть задан закон движения руля высоты в виде зависимости угла γ от времени.

Должно быть известно также состояние атмосферы, в которой совершается полет, т. е. нужно знать скорость w ветра, плотность ρ и вязкость μ воздуха и скорость a звука в функции высоты r. Что касается угла атаки α ракеты, то его заранее задавать нельзя[1]); угол α вычисляется с помощью угла β и воздушной скорости V.

Будем исходить из свойств стандартной атмосферы и зададим типовые характеристики ракеты и ее двигателя. Тогда, если угол отклонения руля высоты задан в функции времени, траектория ракеты может быть определена путем интегрирования системы уравнений (13.14). Действительное проведение вычислений может быть выполнено с помощью электромеханического вычислительного устройства. Эту траекторию типовой ракеты в стандартной атмосфере можно назвать *нормальной траекторией полета*[2]).

Главной характеристикой нормальной траектории служит ее дальность; дальностью называется расстояние между точкой запуска и точкой падения ракеты. Следовательно, задача навигации заключается в определении надлежащего момента времени выключения двигателя ракеты и надлежащего закона перемещения руля в полете так, чтобы осуществить полет на заданную дальность. Эту задачу навигации для полета типовой ракеты в стандартной атмосфере можно решить математически до фактического запуска ракеты, поскольку все данные, необходимые для определения нормальной траектории, известны или заданы заранее.

13.2. Уравнения возмущенного движения. Истинные характеристики атмосферы не обязательно совпадают с характеристиками стандартной атмосферы. Скорость ветра на каждой высоте меняется в зависимости от условий

[1]) Если пренебречь колебаниями ракеты на траектории ($I-0$), то при данных M и Re угол α зависит только от γ и определяется уравнением $C_M=0$. Применяя для обозначения аэродинамических производных символы m_z^α, m_z^γ [24], найдем α в виде $\alpha=-\frac{m_z^\gamma\gamma+m_{z0}}{m_z^\alpha}$; m_{z0}, m_z^α и m_z^γ суть функции M и Re.—*Прим. перев.*

[2]) В русской литературе в этом случае пользуются термином «опорная траектория».—*Прим. перев.*

погоды; температура T воздуха также представляет собой переменную величину. Поэтому благодаря влиянию изменений состояния атмосферы следует ожидать отклонений ракеты от нормальной траектории. Кроме того, данный образец ракеты может несколько отличаться от стандартной ракеты по весу, по свойствам ее двигателя и т. д. Поэтому истинная траектория ракеты будет отличной от нормальной, если отклонить руль ракеты по закону, принятому для нормальной траектории. Таким образом, задача навигации применительно к реальной ракете заключается в корректировании закона перемещения руля таким образом, чтобы дальность в истинном полете совпала с дальностью в полете по нормальной траектории и ракета точно попала в назначенную точку. Ввиду большой скорости полета эту навигационную задачу нельзя решать обычным методом, связанным с полным пренебрежением динамическими эффектами и основанным только на кинематических соотношениях. Вместо этого задачу следует решать с помощью автоматической вычислительной системы, реагирующей практически мгновенно на всякое отклонение от условий нормального полета. Вследствие этого такую задачу целесообразнее называть задачей наведения ракеты, а систему регулирования — системой наведения.

Общая задача наведения в строгом смысле является очень трудной. Однако следует ожидать, что отклонения от нормальных условий будут малы, поскольку нормальная траектория хорошо отражает средние условия полета. Это обстоятельство немедленно наводит на мысль о том, что при оценке возмущений нужно рассматривать лишь величины первого порядка. Такая «линеаризация» и составляет основу баллистической теории возмущений. В результате мы придем к системе, коэффициенты которой определяются их значениями вдоль нормальной траектории и, в общем случае, являются функциями времени. Поэтому в баллистической теории возмущений основную роль играют линейные уравнения с переменными коэффициентами. Таким образом, наше исследование о задаче наведения ракеты дальнего действия служит примером расчета программы управления в таких системах. Конкретное назначение системы определяется здесь требованием уничтожения ошибки по дальности. Регулируемым «входом» служит

поправка на отклонение руля высоты ракеты. Ниже мы разовьем эти положения подробнее.

Условимся обозначать буквами с черточкой соответствующие переменные для нормальной траектории, а отклонения от нее по этим переменным — теми же буквами с символом δ перед ними. Тогда величины переменных в рассматриваемом полете определятся соотношениями

$$\left.\begin{array}{lll} r=\bar{r}+\delta r, & \theta=\bar{\theta}+\delta\theta, & \beta=\bar{\beta}+\delta\beta, \\ v_r=\bar{v}_r+\delta v_r, & v_\theta=\bar{v}_\theta+\delta v_\theta, & \dot{\beta}=\dot{\bar{\beta}}+\delta\dot{\beta}. \end{array}\right\} \qquad (13.15)$$

Отклонения параметров, характеризующих истинное состояние атмосферы, от их значений для стандартной атмосферы определяются отклонениями $\delta\rho$ плотности, δT температуры и δw скорости ветра. Таким образом,

$$\rho=\bar{\rho}+\delta\rho, \qquad T=\bar{T}+\delta T, \qquad w=\bar{w}+\delta w. \qquad (13.16)$$

Если мы предположим, что состав атмосферы на всех высотах не отличается от состава стандартной атмосферы на тех же высотах, то для подсчета отклонений атмосферного давления, если это необходимо, достаточно знать отклонения $\delta\rho$ и δT. Будем считать, что отличие рассматриваемого образца ракеты от стандартного сводится только к погрешности δW в весе и погрешности δI в моменте инерции, так что

$$W=\bar{W}+\delta W; \qquad I=\bar{I}+\delta I. \qquad (13.17)$$

Тягу S мы считаем отвечающей ее стандартной величине. Площадь A крыла и аэродинамические характеристики ракеты, определенные соотношениями (13.10), (13.12), также принимаются неизменными.

Подставив значения переменных, выражаемые равенствами (13.15) — (13.17), в уравнения (13.14) и сократив члены, отвечающие полету по нормальной траектории[1]), мы получим на основании принципа линеаризации сле-

[1]) То есть члены, отвечающие невозмущенному движению (движению по нормальной или опорной траектории). — *Прим. перев.*

дующие уравнения:

$$\left.\begin{aligned}
&\frac{d\,\delta r}{dt} = \delta v_r,\\
&\frac{d\,\delta\theta}{dt} = -\frac{\bar{v}_0}{r^2}\,\delta r + \frac{1}{r}\,\delta v_\theta,\\
&\frac{d\,\delta\beta}{dt} = \delta\dot{\beta};
\end{aligned}\right\} \tag{13.18}$$

$$\left.\begin{aligned}
\frac{d\delta v_r}{dt} &= a_1\delta r + a_2\delta\beta + a_3\delta v_r + a_4\delta v_\theta + a_5\delta\gamma + a_6\delta\rho + \\
&\quad + a_7\delta T + a_8\delta w + a_9\delta W,\\
\frac{d\delta v_\theta}{dt} &= b_1\delta r + b_2\delta\beta + b_3\delta v_r + b_4\delta v_\theta + b_5\delta\gamma + b_6\delta\rho + \\
&\quad + b_7\delta T + b_8\delta w + b_9\delta W,\\
\frac{d\delta\dot{\beta}}{dt} &= c_1\delta r + c_2\delta\beta + c_3\delta v_r + c_4\delta v_\theta + c_5\delta\gamma + c_6\delta\rho + \\
&\quad + c_7\delta T + c_8\delta w + c_9\delta W + c_{10}\delta I.
\end{aligned}\right\} \tag{13.19}$$

Коэффициенты a_i, b_j, c_k суть частные производные от функций F, G и H, определенных в системе (13.14), по соответствующим переменным, вычисленные вдоль траектории. Например,

$$\left.\begin{aligned}
&a_1 = \left(\overline{\frac{\partial F}{\partial r}}\right), \quad a_2 = \left(\overline{\frac{\partial F}{\partial \beta}}\right), \quad a_3 = \left(\overline{\frac{\partial F}{\partial v_r}}\right),\\
&a_4 = \left(\overline{\frac{\partial F}{\partial v_\theta}}\right), \quad a_5 = \left(\overline{\frac{\partial F}{\partial \gamma}}\right), \quad a_6 = \left(\overline{\frac{\partial F}{\partial \rho}}\right),\\
&a_7 = \left(\overline{\frac{\partial F}{\partial T}}\right), \quad a_8 = \left(\overline{\frac{\partial F}{\partial w}}\right), \quad a_9 = \left(\overline{\frac{\partial F}{\partial W}}\right).
\end{aligned}\right\} \tag{13.20}$$

Подробности вычисления этих коэффициентов пояснены в приложении к настоящей главе.

Уравнения (13.18) и (13.19) представляют собой линейные уравнения с переменными коэффициентами относительно шести возмущений. Если известны отклонения $\delta\rho$, δT и δw, характеризующие возмущенное состояние атмосферы, и заданы отклонения $\delta\gamma$, δW, δI, то эта система

дифференциальных уравнений определяет возмущения δr, $\delta\theta$, $\delta\beta$, δv_r, δv_θ и $\delta\dot{\beta}$. Однако, не в разыскании этих возмущений заключается задача наведения. В самом деле, мы ищем функцию $\delta\gamma$ — поправку на изменение угловой координаты руля высоты, требующуюся для уничтожения ошибки по дальности. Как показывает исследование Дреника, эту задачу лучше всего решать методом *сопряженных функций* Блисса[1]).

13.3. Сопряженные функции. Существо метода сопряженных функций заключается в следующем. Пусть функции $y_i(t)$ при $i=1, \ldots, n$ определяются системой n линейных уравнений

$$\frac{dy_i}{dt} - \sum_{j=1}^{n} a_{ij}y_j = Y_i(t), \qquad i=1, \ldots, n, \tag{13.21}$$

где a_{ij} — заданные коэффициенты, которые могут зависеть от времени t. Функции $Y_i(t)$ суть «вынуждающие» функции или входы. Введем совокупность новых функций $\lambda_i(t)$ $(i=1, \ldots, n)$, называемых сопряженными функциями по отношению к функциям $y_i(t)$, которые удовлетворяют следующей системе однородных уравнений:

$$\frac{d\lambda_i}{dt} + \sum_{j=1}^{n} a_{ji}\lambda_j = 0, \qquad i=1, \ldots, n. \tag{13.22}$$

Умножив обе части каждого из уравнений (13.21) на λ_i, а обе части каждого из уравнений (13.22) на y_i и произведя суммирование по индексу i, получим

$$\frac{d}{dt}\sum_{i=1}^{n} \lambda_i y_i - \sum_{i=1}^{n}\sum_{j=1}^{n} (a_{ij}\lambda_i y_j - a_{ji}\lambda_j y_i) = \sum_{i=1}^{n} \lambda_i Y_i.$$

Очевидно, что две части, составляющие двойную сумму, взаимно уничтожаются и

$$\frac{d}{dt}\sum_{i=1}^{n} \lambda_i y_i = \sum_{i=1}^{n} \lambda_i Y_i. \tag{13.23}$$

---

[1]) Bliss G. A., Mathematics for exterior ballistics, New York, 1944.

Проинтегрируем это уравнение по времени в пределах от $t=t_1$ до $t=t_2$:

$$\sum_{i=1}^{n} \lambda_i y_i \Big|_{t=t_2} = \sum_{i=1}^{n} \lambda_i y_i \Big|_{t=t_1} + \int_{t_1}^{t_2} \left(\sum_{i=1}^{n} \lambda_i Y_i \right) dt. \quad (13.24)$$

Блисс называет это соотношение *основной формулой*.

В задаче о полете ракеты дальнего действия величины y_i представляют собой возмущения ($n=6$) и

$$\left.\begin{aligned} y_1 &= \delta r, \quad y_2 = \delta\theta, \quad y_3 = \delta\beta, \\ y_4 &= \delta v_r, \quad y_5 = \delta v_\theta, \quad y_6 = \delta\dot{\beta}. \end{aligned}\right\} \quad (13.25)$$

Тогда, в силу уравнений (13.19), сопряженные функции $\lambda_i(t)$ удовлетворяют следующей системе уравнений:

$$\left.\begin{aligned}
-\frac{d\lambda_1}{dt} &= -\frac{\bar{v}_\theta}{\bar{r}^2}\lambda_2 + a_1\lambda_4 + b_1\lambda_5 + c_1\lambda_6, \\
-\frac{d\lambda_2}{dt} &= 0, \\
-\frac{d\lambda_3}{dt} &= a_2\lambda_4 + b_2\lambda_5 + c_2\lambda_6, \\
-\frac{d\lambda_4}{dt} &= \lambda_1 + a_3\lambda_4 + b_3\lambda_5 + c_3\lambda_6, \\
-\frac{d\lambda_5}{dt} &= \frac{1}{\bar{r}}\lambda_2 + a_4\lambda_4 + b_4\lambda_5 + c_4\lambda_6, \\
-\frac{d\lambda_6}{dt} &= \lambda_3.
\end{aligned}\right\} \quad (13.26)$$

Входные функции Y_i определяются равенствами:

$$Y_1 = Y_2 = Y_3 = 0, \quad (13.27)$$

$$\left.\begin{aligned}
Y_4 &= a_5\delta\gamma + a_6\delta\rho + a_7\delta T + a_8\delta w + a_9\delta W, \\
Y_5 &= b_5\delta\gamma + b_6\delta\rho + b_7\delta T + b_8\delta w + b_9\delta W, \\
Y_6 &= c_5\delta\gamma + c_6\delta\rho + c_7\delta T + c_8\delta w + c_9\delta W + c_{10}\delta I.
\end{aligned}\right\} \quad (13.28)$$

13.4. Стабилизация дальности. Уравнения (13.26) не определяют функций $\lambda_1, \ldots, \lambda_6$ полностью. Чтобы это сделать, необходимо задать значения функций $\lambda_1, \ldots, \lambda_6$

в некоторый момент времени. Какие значения функций $\lambda_1, \ldots, \lambda_6$ задать и в какой момент времени – это зависит от конкретного назначения системы регулирования. В нашей задаче о наведении ракеты на цель мы хотим уничтожить ошибку по дальности. Поэтому нас интересует величина $\delta\theta$, измеренная в момент удара ракеты о землю, т. е. величина $\delta\theta_2$, где индекс 2 относится к значению данной переменной в конце траектории. Этого достаточно, как мы сейчас увидим, для полного определения функций $\lambda_1, \ldots, \lambda_6$.

Пусть t_2 есть момент достижения ракетой конечной точки траектории в ее истинном полете, а $\bar{t}_2$ – момент достижения ракетой конечной точки при полете по нормальной траектории. Тогда

$$t_2 = \bar{t}_2 + \delta t_2. \tag{13.29}$$

Подобным же образом

$$\left.\begin{aligned} r_2 &= \bar{r}_2 + \delta r_2, \\ \theta_2 &= \bar{\theta}_2 + \delta\theta_2. \end{aligned}\right\} \tag{13.30}$$

Но

$$\left.\begin{aligned} \delta r_2 &= (\bar{v}_r)_{t=\bar{t}_2}\,\delta t_2 + (\delta r)_{t=\bar{t}_2}, \\ \delta\theta_2 &= \frac{1}{r_0}\,(\bar{v}_\theta)_{t=\bar{t}_2}\,\delta t_2 + (\delta\theta)_{t=\bar{t}_2}. \end{aligned}\right\} \tag{13.31}$$

Однако величина δr_2 по определению равна нулю, ибо под конечной точкой траектории понимается точка, в которой ракета касается земной поверхности, т. е. $r_2 = \bar{r}_2 = r_0$. Исключив δt_2 из уравнений (13.31), получим

$$\delta\theta_2 = \left[-\frac{1}{\bar{r}}\left(\frac{\bar{v}_\theta}{\bar{v}_r}\right)\delta r + \delta\theta \right]_{t=\bar{t}_2}. \tag{13.32}$$

Поэтому, задав значения функций $\lambda_1, \ldots, \lambda_6$ в момент $t = \bar{t}_2$ соприкосновения ракеты с землей, с помощью соотношений

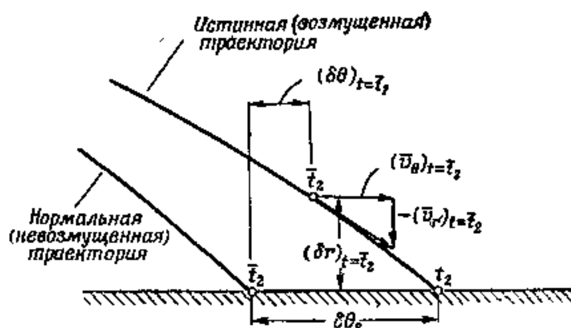
$$\left.\begin{aligned} &\lambda_1 = -\frac{1}{\bar{r}}\left(\frac{\bar{v}_\theta}{\bar{v}_r}\right), \qquad \lambda_2 = 1, \\ &\lambda_3 = \lambda_4 = \lambda_5 = \lambda_6 = 0 \end{aligned}\right\} \tag{13.33}$$

определим ошибку по дальности посредством равенства

$$\delta\theta_2 = \sum_{i=1}^{n} \lambda_i y_i \Big|_{t=\bar{t}_2} =$$

$$= [\lambda_1 \delta r + \lambda_2 \delta\theta + \lambda_3 \delta\beta + \lambda_4 \delta v_r + \lambda_5 \delta v_\theta + \lambda_6 \delta\dot{\beta}]_{t=\bar{t}_2}. \quad (13.34)$$

Если нормальная траектория найдена, то коэффициенты уравнений (13.27) оказываются известными функциями времени. Тогда эти уравнения вместе с граничными значениями в конце траектории определят сопряженные функции $\lambda_1, \ldots, \lambda_6$. Интегрирование следует вести в «обратном»



Фиг. 116

направлении при $t < \bar{t}_2$, осуществляя его, если это возможно, с помощью электромеханического интегратора. Когда сопряженные функции будут определены, можно применить основную формулу (13.25) и преобразовать выражение (13.34) ошибки по дальности следующим образом: пусть $\bar{t}_1$ означает момент времени, в который производится выключение двигателя при полете по нормальной траектории. Тогда условие уничтожения ошибки $\delta\theta_2$ по дальности можно выразить с помощью соотношения

$$\delta\theta_2 = 0 = [\lambda_1 \delta r + \lambda_2 \delta\theta + \lambda_3 \delta\beta + \lambda_4 \delta v_r + \lambda_5 \delta v_\theta + \lambda_6 \delta\dot{\beta}]_{t=\bar{t}_1} +$$

$$+ \int\limits_{t=\bar{t}_1}^{t=\bar{t}_2} (\lambda_4 Y_4 + \lambda_5 Y_5 + \lambda_6 Y_6)\, dt. \quad (13.35)$$

Это и есть основное уравнение задачи наведения; мы используем его в последующих пунктах.

13.5. Определение момента выключения двигателя. Условие (13.35), справедливое при произвольных возмущениях, можно расчленить на две части и приравнять нулю сумму и интеграл, фигурирующие в этом условии, независимо. Поэтому в момент $\bar{t}_1$ отключения двигателя на нормальной траектории имеет место равенство

$$[\lambda_1\delta r+\lambda_2\delta\theta+\lambda_3\delta\beta+\lambda_4\delta v_r+\lambda_5\delta v_\theta+\lambda_6\delta\dot{\beta}]_{t=\bar{t}_1}=0. \quad (13.36)$$

Так как момент $\bar{t}_1$ отключения двигателя определяет стандартный момент времени, не обязательно совпадающий с действительным моментом t_1 отключения двигателя, т. е.

$$t_1=\bar{t}_1+\delta t_1, \quad (13.37)$$

то соотношение (13.36) следует привести к более целесообразной форме, характеризующей истинный момент отключения. Это легко сделать, так как с точностью до величины первого порядка имеем на основании равенства (13.35)

$$(\delta r)_{t=\bar{t}_1}=(r)_{t=t_1}-\left(\overline{\frac{dr}{dt}}\right)_{t=\bar{t}_1}\delta t_1-(\bar{r})_{t=\bar{t}_1}$$

или

$$(\delta r)_{t=\bar{t}_1}=(r)_{t=t_1}-(\bar{r})_{t=\bar{t}_1}-(\bar{v}_r)_{t=\bar{t}_1}\delta t_1.$$

Подобным же образом

$$(\delta\theta)_{t=\bar{t}_1}=(\theta)_{t=t_1}-(\bar{\theta})_{t=\bar{t}_1}-\left(\frac{1}{\bar{r}}\bar{v}_\theta\right)_{t=\bar{t}_1}\delta t_1,$$

$$(\delta\beta)_{t=\bar{t}_1}=(\beta)_{t=t_1}-(\bar{\beta})_{t=\bar{t}_1}-(\bar{\dot{\beta}})_{t=\bar{t}_1}\delta t_1,$$

$$(\delta v_r)_{t=\bar{t}_1}=(v_r)_{t=t_1}-(\bar{v}_r)_{t=\bar{t}_1}-(\bar{F})_{t=\bar{t}_1}\delta t_1,$$

$$(\delta v_\theta)_{t=\bar{t}_1}=(v_\theta)_{t=t_1}-(\bar{v}_\theta)_{t=\bar{t}_1}-(\bar{G})_{t=\bar{t}_1}\delta t_1,$$

$$(\delta\dot{\beta})_{t=\bar{t}_1}=(\dot{\beta})_{t=t_1}-(\bar{\dot{\beta}})_{t=\bar{t}_1}-(\bar{H})_{t=\bar{t}_1}\delta t_1,$$

где $\overline{F}$, $\overline{G}$ и $\overline{H}$ суть значения величин F, G, H, определенных уравнениями (13.14), вдоль нормальной траектории. В действительности эти величины следует брать в некоторый момент времени, непосредственно предшествующий моменту $\overline{t}_1$ выключения двигателя на нормальной траектории, что необходимо для учета влияния сил, обусловливающих ускорение ракеты, ибо это ускорение определяется также и тягой.

Определим теперь величины J и $\overline{J}$ равенствами

$$J = [\lambda_1^* r + \lambda_2^* \theta + \lambda_3^* \beta + \lambda_4^* v_r + \lambda_5^* v_\theta + \lambda_6^* \dot{\beta}]_{t=t_1} \qquad (13.38)$$

и

$$\overline{J} = [\lambda_1^* \overline{r} + \lambda_2^* \overline{\theta} + \lambda_3^* \overline{\beta} + \lambda_4^* \overline{v}_r + \lambda_5^* \overline{v}_\theta + \lambda_6^* \overline{\dot{\beta}}]_{t=\overline{t}_1},$$

где λ_i^* суть значения функций λ_i в момент $\overline{t}_1$ выключения двигателя на нормальной траектории. Тогда в момент t_1 выключения двигателя на истинной траектории должно выполняться условие

$$J = \overline{J} + \left[\lambda_1^* \overline{v}_r + \lambda_2^* \frac{\overline{v}_\theta}{\overline{r}} + \lambda_3^* \overline{\dot{\beta}} + \lambda_4^* \overline{F} + \lambda_5^* \overline{G} + \lambda_6^* \overline{H} \right]_{t=\overline{t}_1} (t_1 - \overline{t}_1). \qquad (13.39)$$

Таково уравнение, определяющее надлежащий момент выключения двигателя.

Когда нормальная траектория известна, величина $\overline{J}$ и величина, заключенная в квадратные скобки в правой части равенства (13.39), становятся фиксированными. В этих условиях всю правую часть равенства (13.39) можно рассматривать как линейную возрастающую функцию времени t, если только вместо t_1 подставить t. Одновременно с помощью величин λ_i^*, вычисленных заранее, и значений координат и скоростей ракеты в фактическом полете, измеряемых на наблюдательных станциях, можно определять значения J в каждый момент времени, предшествующий моменту выключения двигателя. В результате мы получаем возможность непрерывно сравнивать величины, стоящие в обеих частях равенства (13.39). Когда эти величины сравняются, равенство (13.39) удовлетворится. Тогда посылается сигнал выключения двигателя и горение в ракете прекращается.

13.6. Условия попадания ракеты в цель. Когда ракетный двигатель выключается раньше или позже стандартного момента $\bar{t}_1$ выключения, то благодаря влиянию массы горючего, оставшегося в баке (если только бак к этому моменту не опустошен), величины веса W и момента I инерции ракеты принимают значения, отличные от соответствующих значений этих величин для нормальной траектории. Не исключено также, что полезный груз ракеты не совпадает с полезным грузом, соответствующим стандартной ракете. Тогда, по выключении двигателя, возникают постоянные отклонения δW и δI, постоянные в том смысле, что они не меняются во времени и становятся известными сразу же после выключения двигателя. Иной характер имеют отклонения $\delta\rho$, δT и δw истинной атмосферы от стандартной. Их мы не можем узнать иначе, как путем измерения. В дальнейшем мы ставим своей задачей применить в качестве измерительного инструмента саму ракету, для чего поступим следующим образом.

После того как удовлетворено условие выключения двигателя, условие уничтожения ошибки по дальности заключается в обращении в нуль интеграла в правой части равенства (13.35). Поскольку Y_i в подинтегральной функции зависят от произвольных возмущений $\delta\rho$, δT и δw, заранее неизвестных, то последнее условие может быть выполнено только при обращении в нуль самого подинтегрального выражения[1]). Следовательно, в силу соотношений (13.28), получим

$$\begin{aligned}(\lambda_4 a_5 &+ \lambda_5 b_5 + \lambda_6 c_5)\,\delta\gamma + \\ &+ (\lambda_4 a_6 + \lambda_5 b_6 + \lambda_6 c_6)\,\delta\rho + (\lambda_4 a_7 + \lambda_5 b_7 + \lambda_6 c_7)\,\delta T + \\ &+ (\lambda_4 a_8 + \lambda_5 b_8 + \lambda_6 c_8)\,\delta w + (\lambda_4 a_9 + \lambda_5 b_9 + \lambda_6 c_9)\,\delta W + \\ &\qquad + \lambda_6 c_{10}\,\delta I = 0;\end{aligned}$$

---

[1]) Чтобы $\int\limits_{t_1}^{t_2} F(\tau)\,\delta\rho\cdot d\tau$ обращался в нуль при *любых* непрерывных функциях $\delta\rho$, необходимо, чтобы было $F \equiv 0$. На этом свойстве основан, в частности, вывод уравнения Лагранжа в механике и вариационном исчислении.—*Прим. перев.*

введя обозначения

$$\left.\begin{aligned} d_5 &= \lambda_4 a_5 + \lambda_5 b_5 + \lambda_6 c_5, \\ d_6 &= \lambda_4 a_6 + \lambda_5 b_6 + \lambda_6 c_6, \\ d_7 &= \lambda_4 a_7 + \lambda_5 b_7 + \lambda_6 c_7, \\ d_8 &= \lambda_4 a_8 + \lambda_5 b_8 + \lambda_6 c_8, \\ D &= -(\lambda_4 a_9 + \lambda_5 b_9 + \lambda_6 c_9)\,\delta W - \lambda_6 c_{10}\,\delta I, \end{aligned}\right\} \tag{13.40}$$

запишем это условие в виде

$$d_5\,\delta\gamma + d_6\,\delta\rho + d_7\,\delta T + d_8\,\delta w = D. \tag{13.41}$$

Уравнения (13.19) можно переписать в форме

$$\left.\begin{aligned} a_5\,\delta\gamma + a_6\,\delta\rho + a_7\,\delta T + a_8\,\delta w &= A, \\ b_5\,\delta\gamma + b_6\,\delta\rho + b_7\,\delta T + b_8\,\delta w &= B, \\ c_5\,\delta\gamma + c_6\,\delta\rho + c_7\,\delta T + c_8\,\delta w &= C, \end{aligned}\right\} \tag{13.42}$$

где

$$\left.\begin{aligned} A &= \frac{d\,\delta v_r}{dt} - a_1\,\delta r - a_2\,\delta\beta - a_3\,\delta v_r - \\ &\qquad - a_4\,\delta v_\theta - a_9\,\delta W, \\ B &= \frac{d\,\delta v_\theta}{dt} - b_1\,\delta r - b_2\,\delta\beta - b_3\,\delta v_r - \\ &\qquad - b_4\,\delta v_\theta - b_9\,\delta W, \\ C &= \frac{d\delta\dot{\beta}}{dt} - c_1\,\delta r - c_2\,\delta\beta - c_3\,\delta v_r - \\ &\qquad - c_4\,\delta v_\theta - c_9\,\delta W - c_{10}\,\delta I. \end{aligned}\right\} \tag{13.43}$$

Если на станциях, с которых производится измерение параметров летящей ракеты, измерить величины A, B и C и если, кроме того, одно из трех возмущений $\delta\rho$, δT и δw измерить независимо при помощи прибора, находящегося внутри ракеты, то, пользуясь этими результатами и любыми двумя из уравнений (13.42), можно будет определить остальные два атмосферных возмущения. Это, по существу, сводится к использованию для определения возмущений $\delta\rho$, δT и δw самой ракеты в качестве измерительного прибора. Следует заметить, что при определении всех трех

возмущений $\delta\rho$, δT и δw мы не можем ограничиться только уравнениями (13.42), не пользуясь добавочным прибором, так как мы не знаем $\delta\gamma$ и уравнения (13.41) и (13.42) не независимы.

Если, например, возмущение δT определено независимо, при помощи прибора, то уравнения (13.41) и (13.42) нужно переписать в виде

$$\begin{aligned} a_5\ \delta\gamma + a_6\ \delta\rho + a_8\ \delta w &= A - a_7\ \delta T,\\ b_5\ \delta\gamma + b_6\ \delta\rho + b_8\ \delta w &= B - b_7\ \delta T,\\ d_5\ \delta\gamma + d_6\ \delta\rho + d_8\ \delta w &= D - d_7\ \delta T. \end{aligned}$$

Из этих уравнений можно определить необходимую поправку на отклонение руля в каждый момент времени в функции значений величин a_i, b_j, d_k и A, B, D для того же момента времени. Последние величины частично зависят от параметров нормальной траектории, определенных заранее, а частично – от данных, полученных с помощью телеметрических измерений координат и скоростей ракеты в полете. На больших высотах, где плотность воздуха весьма мала, аэродинамические силы почти пренебрежимы по сравнению с силой тяжести и силами инерции. В этом случае величины A, B и C, определяемые уравнениями (13.43), представляют собой малые разности больших величин и поэтому точное определение этих трех величин сопряжено с весьма большими затруднениями. Если руль высоты летящей ракеты отклонять в соответствии с программой, определяемой соотношением (13.44), и произвести выключение двигателя в надлежащий момент времени, найденный в предыдущем пункте, то ракета будет наводиться на выбранную цель на земной поверхности, несмотря на атмосферные возмущения.

13.7. Система наведения ракеты. Как только на основании технического задания на полет ракеты установлен общий характер ее траектории, первая задача заключается в определении нормальной траектории (в стандартной атмосфере) по ожидаемым значениям параметров ракеты при ее стандартном весе. Зная нормальную траекторию, мы можем найти и функции a_i, b_j и c_k. С помощью уравнения (13.27) и условий (13.33) на конце траектории можно определить сопряженные функции λ_i. Всеми этими

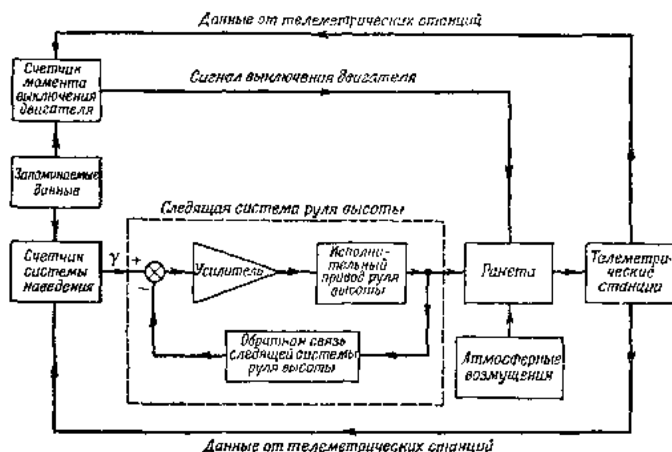
данными необходимо располагать до начала реального полета ракеты, поэтому их можно назвать «запоминаемыми данными».

До момента выключения двигателя рулем высоты можно управлять в соответствии с программой, составленной для полета по нормальной траектории, а стабилизация ракеты на этой траектории осуществляется с помощью дополнительных реактивных сопел или вспомогательных ракет. Однако телеметрические станции вступают в действие немедленно после запуска ракеты и подают на ее приемные устройства данные о ее положении в пространстве и скоростях полета. Эти данные поступают, прежде всего, на счетчик момента выключения двигателя, который, используя запоминаемые данные, непрерывно производит сравнение величин, находящихся в обеих частях соотношения (13.39), выражающего условие выключения двигателя. Как только это условие будет удовлетворено, произойдет выключение двигателя.

В момент выключения двигателя сигналы с телеметрических станций, до той поры подававшиеся на счетчик этого выключателя, переключаются на вычислительное устройство системы наведения. Продолжительность протекшего времени от момента запуска до момента выключения определяет также количество горючего, оставшегося в баках, и тем самым определяет отклонения веса и момента инерции ракеты от стандартных значений, соответствующих моменту выключения двигателя. Тогда на основании этих данных, вместе с запоминаемыми данными полета по нормальной траектории, вычислительное устройство может вырабатывать поправку $\delta\gamma$ на отклонение руля в соответствии с соотношениями (13.40), (13.43) и (13.44). Теоретически величина $\delta\gamma$ должна быть найдена без запаздывания по отношению к моменту времени получения входных данных, так как равенство (13.44) представляет собой условие равенства двух величин, рассматриваемых в одинаковые моменты времени. В результате истинный угол γ отклонения руля высоты в фактическом полете определяется с помощью вычисленной поправки $\delta\bar{\gamma}$ в сочетании с углом $\bar{\gamma}$ отклонения этого руля, характеризующим полет по нормальной траектории. Начиная с этого момента процесс разработки механизма управления рулем высоты можно вести на основании

методики, предназначенной для разработки обычных следящих систем (с обратной связью), применяя обычные критерии быстродействия, устойчивости и точности отработки. В свете этих рассуждений общую схему системы наведения можно представить в форме, показанной на фиг. 117.

Вычислительные устройства, предусмотренные в этой схеме, находятся на борту ракеты и получают данные о ее положении и скоростях от неподвижных наземных наблюдательных станций, расположенных вдоль траектории. Эта



Фиг. 117

передача данных и образует цепь обратной связи, как показано на фиг. 117. Если вычислительные устройства выполнены надлежащим образом, то их работа обеспечивает заданное качество системы; это означает, что действие этих вычислительных устройств равносильно действию корректирующего контура или усилителя обычной следящей системы. Следовательно, в широком плане наша система наведения весьма близка к простым следящим системам, рассмотренным в предыдущих главах. Однако система наведения все же очень сложна. Ее разработка требует применения теории возмущений совместно с теорией сопряженных функций.

Хотя разобранный нами пример управления полетом ракеты дальнего действия изложен, пожалуй, в слишком упрощенной форме, он служит наглядной иллюстрацией методики приложения теории к разработке систем автоматического управления. В этом примере мы исходили только из одного критерия проектирования, а именно требовали уничтожения ошибки по дальности. В более сложных системах можно сформулировать несколько критериев проектирования и может возникнуть необходимость прибегнуть к использованию нескольких систем сопряженных функций. Тем не менее принцип решения задачи останется тем же, что и в данном простом примере.

13. 8. Вычислительные устройства, применяемые в системах автоматического регулирования. Хотя в нашу задачу не входит описание подробностей конструкции и технических особенностей тех или иных узлов, входящих в состав рассмотренных выше систем, все же целесообразно охарактеризовать здесь в общей форме свойства вычислительных устройств и требования к ним, учитывая особую важность этих устройств в наиболее современных системах регулирования, отмеченную в первый раз в связи с задачей об оптимальном законе переключения (гл. X). Читателю же, интересующемуся этим вопросом более подробно, следует обратиться к специальным руководствам по этой теме[1] [2].

В настоящее время распространены два вида вычислительных устройств: физические модели и цифровые счетные машины. Физической моделью, как это точно следует из ее названия, называется устройство, физически моделирующее задачу того типа, которую проектировщик модели желает решать. Следовательно, физическая модель представляет собой систему, которая описывается с помощью тех же математических операций, с помощью которых описывается и явление, подлежащее расчету.

Входные величины подаются на это устройство в виде значений некоторой физической величины—электрического

---

[1]) Engineering Research Associates, «High-speed Computing Devices», New York, 1950. По электромоделям см. Корн Г. и Корн Т., Электронные моделирующие устройства, М., 1955.

[2]) См. [5], т. III, и [16].—*Прим. перев.*

напряжения или тока, углового перемещения некоторого вала или упругой силы пружины. Устройство перерабатывает эти входы в другие физические величины—выходы—в соответствии со своей структурой, выбранной его проектировщиком так, чтобы процессы в модели характеризовали определенные математические преобразования. Таким образом, входы устройства-модели представляют собой приспособления, считывающие различные физические величины системы автоматического регулирования, а ее выходы служат управляющими сигналами, которые непосредственно поступают на индивидуальные регуляторы регулируемых переменных.

В отличие от моделирующих устройств цифровая машина работает, как счетчик. Входные данные задачи задаются с помощью цифр; машина перерабатывает эти данные в соответствии с арифметическими правилами или другой системой правил, обусловливаемой существом задачи, решаемой в цифровом виде, и выражает окончательный результат в числовой форме. Этот принцип действия цифровой машины влечет за собой два очень важных следствия. Во-первых, конструкция устройства для ввода входных данных и выдачи выходных результатов, т. е. «преобразователей», должна быть приспособлена для осуществления надлежащего перевода операций, производимых машиной в соответствии с заданной ей логикой, в физические величины регулируемой системы и наоборот. Во-вторых, задача, подлежащая решению, должна быть подготовлена в виде, пригодном для ввода ее в цифровую машину. В случае моделирующего устройства такая задача, по существу, уже подготовлена в силу самой структуры модели; в цифровой же машине конструкция определяется не какой-либо конкретной задачей или классом задач, а системой логических операций, в соответствии с которой машина и должна действовать при решении конкретной задачи.

Сравнивая моделирующие устройства и цифровые машины, как звенья в системах автоматического регулирования, мы заметим, прежде всего, что в простых прикладных задачах моделирующее устройство почти всегда оказывается менее сложным по сравнению с цифровой машиной, предназначенной для той же цели. Даже самая элементарная цифровая машина требует блоков выполнения арифмети-

ческих операций, ячеек памяти, регулирующих цепей и устройств для ввода входных данных и выдачи выходных данных. При решении простых задач этот набор приспособлений оказывается непроизводительно сложным. Наоборот, моделирующее устройство нет надобности делать более сложным, чем этого требует задача. В самом деле, как уже отмечалось выше, корректирующий контур в простой следящей системе уже представляет собой такое моделирующее устройство.

В задачах же, в которых вычислительные операции оказываются более сложными (как в задаче о наведении ракеты на цель, разобранной в этой главе), моделирующие устройства утрачивают свои преимущества, и мы обнаруживаем здесь второе коренное различие между машинами двух типов. Моделирующее устройство, представляющее собой физическую аналогию рассматриваемой задачи, должно быть более сложным для более сложных моделируемых задач. Если модель механическая, то в нее приходится вводить все более и более длинные цепи механических передач, фрикционных интеграторов и других узлов, соединяемых вместе; если модель электрическая, то приходится вводить все большее и большее число каскадов усиления. В случае механической модели неизбежные зазоры в передачах и сочленениях, хотя и находящиеся в пределах допусков в простых цепях, будут с увеличением числа звеньев возрастать до такой степени, что в конце концов суммарный зазор или «игра» в машине превзойдет заметные значения выходной переменной и устройство станет бесполезным. В случае электрической модели подобным же образом будут накапливаться случайные электрические возмущения и шумы, всегда присутствующие в электрических цепях, величина которых в конце концов превзойдет полезный сигнал. Поскольку электрический шум гораздо менее вреден, чем зазоры в механических передачах, электрические модели можно делать более сложными по сравнению с их механическими аналогами, но есть предел и для них. Наоборот, цифровая машина совершенно свободна от таких случайных явлений, как зазоры или шумы, и для нее не существует ограничений на степень сложности задач, которые такая машина может обработать.

Третье важное различие между моделирующими устрой-

ствами и цифровыми машинами связано со степенью точности вычислений, заложенной в каждой из этих систем. Степень точности моделирующего устройства ограничена степенью точности выполнения преобразований над физическими величинами и измерения этих величин и степенью точности, с которой физическая система описывается идеализированными законами. На практике самая лучшая модель может достичь точности в 0,01%; многие модели имеют точность только в пределах 1—2%. Для некоторых приложений такая степень точности терпима, в других случаях она недостаточна. Цифровая же машина в отличие от модели оперирует только с числами, и ее можно сделать сколь угодно точной. Для повышения точности мы должны лишь увеличить число значащих цифр, вырабатываемых машиной для определения каждой из величин. Конечно, точность всей машины попрежнему ограничена точностью преобразователей входных и выходных величин, но это замечание не изменяет того обстоятельства, что в задачах, требующих высокой точности, цифровая машина предпочтительнее моделирующего устройства.

Эти два вида вычислительных устройств различаются еще в одном отношении, четвертом по счету. Физическая модель работает, как говорят, в естественном времени. Это означает, что модель непрерывно вырабатывает решение задачи, которую она решает, и что это решение в каждый момент времени соответствует всем входным величинам, введенным в модель до данного момента[1]). Что же касается цифровой машины, то она действует, отрабатывая и решая определенную программу, составленную в соответствии со счетными операциями задачи. Поэтому цифровая машина выдает результат вычислений лишь в виде последовательности значений выходной величины в дискретные моменты времени; кроме того, поскольку продолжительность выполнения вычислений в машине является хотя и малой, но конечной, выход запаздывает по отношению к входу. В связи с этим возникают две задачи: задача интерполяции выходной величины в промежутках между дискрет-

[1]) Существуют также модели, работающие в ускоренном темпе, с многократным повторением решения. См. Гутенмахер Л. И., Электрические модели, М.—Л., 1950.—*Прим. ред.*

ными моментами выдачи выхода и задача о предварении (упреждении) значений выходной величины по ее предыдущим значениям, с тем чтобы устранить ее запаздывание. Разумеется, если продолжительность выполнения вычислений намного меньше характеристического времени регулируемой системы, то вопросом о предварении можно не заниматься и считать, что цифровая машина работает в естественном времени. Современные электронные цифровые машины, повидимому, обладают достаточным быстродействием в этом смысле для использования их в рассматривавшихся выше задачах управления ракетами дальнего действия. Однако в случае управляемых снарядов большой скорости влияние этого запаздывания необходимо должным образом учесть при разработке системы автоматического управления.

Приложение к главе XIII

ПОДСЧЕТ КОЭФФИЦИЕНТОВ В УРАВНЕНИЯХ ВОЗМУЩЕННОГО ДВИЖЕНИЯ

Величины F, G и H введены с помощью уравнений (13.14); они зависят от параметров Σ, Λ, Δ и N. В силу определения этих параметров равенствами (13.2) и (13.13) их можно записать в виде

$$\left.\begin{aligned}\Sigma &= \frac{Sg}{W},\\ \Lambda &= \frac{g}{W}\,\frac{1}{2}\,\rho A C_L \sqrt{v_r^2+(v_\theta+w)^2},\\ \Delta &= \frac{g}{W}\,\frac{1}{2}\,\rho A C_D \sqrt{v_r^2+(v_\theta+w)^2},\\ N &= \frac{1}{I}\,\frac{1}{2}\,\rho A C_M\,[v_r^2+(v_\theta+w)^2],\end{aligned}\right\} \tag{13.45}$$

где аэродинамические коэффициенты C_L, C_D и C_M зависят от угла атаки α, угловой координаты γ руля высоты, числа Маха M и числа Рейнольдса Re. Эти аэродинамические переменные непосредственно выражаются через переменные полета:

$$\alpha=\beta-\operatorname{arctg}\frac{v_r}{v_\theta+w}; \qquad \mathrm{M}=\frac{V}{a(r)}; \qquad \mathrm{Re}=\frac{\rho V l}{\mu(r)}. \tag{13.46}$$

Здесь $a(r)$ — скорость звука в воздухе, $\mu(r)$ — коэффициент вязкости воздуха; обе эти величины суть функции высоты r. В последующих вычислениях тяга S принимается зависящей только от высоты[1]). Мы допускаем также, что состав атмосферы на каждой данной высоте не меняется по сравнению с составом стандартной атмосферы на этих же высотах; меняются лишь плотность ρ и температура T. Таким образом, изменения величин a и μ на каждой данной высоте обусловливаются колебаниями температуры T.

Для Σ имеем

$$\left.\begin{aligned} \frac{\partial \Sigma}{\partial r} &= \frac{g}{W}\frac{\partial S}{\partial r}; \\ \frac{\partial \Sigma}{\partial W} &= -\frac{\Sigma}{W}. \end{aligned}\right\} \tag{13.47}$$

Все остальные частные производные от Σ равны нулю. Дифференцируя Λ, получаем

$$\left.\begin{aligned} \frac{\partial \Lambda}{\partial r} &= \Lambda \left\{ \frac{1}{\rho}\frac{d\rho}{dr}\left(1 + \frac{\mathrm{Re}}{C_L}\frac{\partial C_L}{\partial \mathrm{Re}}\right) + \right. \\ &\quad + \frac{1}{V^2}\frac{dw}{dr}\left[\left(\frac{\mathrm{M}}{C_L}\frac{\partial C_L}{\partial \mathrm{M}} + \frac{\mathrm{Re}}{C_L}\frac{\partial C_L}{\partial \mathrm{Re}} + 1\right)(v_\theta + w) + \right. \\ &\quad\quad \left. + \frac{1}{C_L}\frac{\partial C_L}{\partial \alpha} v_r\right] - \frac{\mathrm{M}}{C_L}\frac{\partial C_L}{\partial \mathrm{M}}\frac{1}{a}\frac{da}{dr} - \\ &\quad\quad\quad \left. - \frac{\mathrm{Re}}{C_L}\frac{\partial C_L}{\partial \mathrm{Re}}\frac{1}{\mu}\frac{d\mu}{dr}\right\}, \\ \frac{\partial \Lambda}{\partial v_r} &= \Lambda \frac{v_r}{V^2}\left(\frac{\mathrm{M}}{C_L}\frac{\partial C_L}{\partial \mathrm{M}} + \frac{\mathrm{Re}}{C_L}\frac{\partial C_L}{\partial \mathrm{Re}} + \right. \\ &\quad\quad \left. + 1 - \frac{1}{C_L}\frac{\partial C_L}{\partial \alpha}\frac{v_\theta + w}{v_r}\right), \\ \frac{\partial \Lambda}{\partial v_\theta} &= \Lambda \frac{v_\theta + w}{V^2}\left(\frac{\mathrm{M}}{C_L}\frac{\partial C_L}{\partial \mathrm{M}} + \right. \\ &\quad\quad \left. + \frac{\mathrm{Re}}{C_L}\frac{\partial C_L}{\partial \mathrm{Re}} + 1 + \frac{1}{C_L}\frac{\partial C_L}{\partial \alpha}\frac{v_r}{v_\theta + w}\right), \\ \frac{\partial \Lambda}{\partial \beta} &= \Lambda \frac{1}{C_L}\frac{\partial C_L}{\partial \alpha}, \end{aligned}\right\} \tag{13.48}$$

---

[1]) Так как при постоянной скорости горения истечение газовой струи из сопла зависит только от противодавления в окрестности сопла, т. е. от атмосферного давления, зависящего, по предположению, только от высоты. — *Прим. перев.*

$$
\left.
\begin{aligned}
&\frac{\partial\Lambda}{\partial\gamma}=\Lambda\frac{1}{C_L}\frac{\partial C_L}{\partial\gamma},\\
&\frac{\partial\Lambda}{\partial\rho}=\Lambda\frac{1}{\rho}\left(1+\frac{\mathrm{Re}}{C_L}\frac{\partial C_L}{\partial\,\mathrm{Re}}\right),\\
&\frac{\partial\Lambda}{\partial T}=-\Lambda\left(\frac{\mathrm{M}}{C_L}\frac{\partial C_L}{\partial\,\mathrm{M}}\frac{1}{2T}+\frac{\mathrm{Re}}{C_L}\frac{\partial C_L}{\partial\,\mathrm{Re}}\frac{1}{\mu}\frac{\partial\mu}{\partial T}\right),\\
&\frac{\partial\Lambda}{\partial w}=\Lambda\frac{v_\theta+w}{V^2}\left(\frac{\mathrm{M}}{C_L}\frac{\partial C_L}{\partial\,\mathrm{M}}+\frac{\mathrm{Re}}{C_L}\frac{\partial C_L}{\partial\,\mathrm{Re}}+\right.\\
&\qquad\qquad\left.+1+\frac{1}{C_L}\frac{\partial C_L}{\partial\alpha}\frac{v_r}{v_\theta+w}\right)=\frac{\partial\Lambda}{\partial v_\theta},\\
&\frac{\partial\Lambda}{\partial W}=-\frac{\Lambda}{W}.
\end{aligned}
\right\}
$$

Частные производные от Δ найдутся с помощью предыдущих соотношений путем подстановки в них величины Δ вместо Λ и величины C_D вместо C_L.

Дифференцируя N, получаем

$$
\left.
\begin{aligned}
&\frac{\partial N}{\partial r}=N\left\{\frac{1}{\rho}\frac{d\rho}{dr}\left(1+\frac{\mathrm{Re}}{C_\mathrm{M}}\frac{\partial C_\mathrm{M}}{\partial\,\mathrm{Re}}\right)+\frac{1}{V^2}\frac{dw}{dr}\left[\left(\frac{\mathrm{M}}{C_\mathrm{M}}\frac{\partial C_\mathrm{M}}{\partial \mathrm{M}}+\right.\right.\right.\\
&\qquad\left.\left.+\frac{\mathrm{Re}}{C_\mathrm{M}}\frac{\partial C_\mathrm{M}}{\partial\,\mathrm{Re}}+2\right)(v_\theta+w)+\frac{1}{C_\mathrm{M}}\frac{\partial C_\mathrm{M}}{\partial\alpha}v_r\right]-\\
&\qquad\left.-\frac{\mathrm{M}}{C_\mathrm{M}}\frac{\partial C_\mathrm{M}}{\partial \mathrm{M}}\frac{1}{a}\frac{da}{dr}-\frac{\mathrm{Re}}{C_\mathrm{M}}\frac{\partial C_\mathrm{M}}{\partial\,\mathrm{Re}}\frac{1}{\mu}\frac{d\mu}{dr}\right\},\\
&\frac{\partial N}{\partial v_r}=N\frac{v_r}{V^2}\left(\frac{\mathrm{M}}{C_\mathrm{M}}\frac{\partial C_\mathrm{M}}{\partial \mathrm{M}}+\frac{\mathrm{Re}}{C_\mathrm{M}}\frac{\partial C_\mathrm{M}}{\partial\,\mathrm{Re}}+2-\frac{1}{C_\mathrm{M}}\frac{\partial C_\mathrm{M}}{\partial\alpha}\frac{v_\theta+w}{v_r}\right),\\
&\frac{\partial N}{\partial v_\theta}=N\frac{v_\theta+w}{V^2}\left(\frac{\mathrm{M}}{C_\mathrm{M}}\frac{\partial C_\mathrm{M}}{\partial \mathrm{M}}+\frac{\mathrm{Re}}{C_\mathrm{M}}\frac{\partial C_\mathrm{M}}{\partial\,\mathrm{Re}}+2+\frac{1}{C_\mathrm{M}}\frac{\partial C_\mathrm{M}}{\partial\alpha}\frac{v_r}{v_\theta+w}\right),\\
&\frac{\partial N}{\partial\beta}=N\frac{1}{C_\mathrm{M}}\frac{\partial C_\mathrm{M}}{\partial\alpha},\\
&\frac{\partial N}{\partial\gamma}=N\frac{1}{C_\mathrm{M}}\frac{\partial C_\mathrm{M}}{\partial\gamma},\\
&\frac{\partial N}{\partial\rho}=N\frac{1}{\rho}\left(1+\frac{\mathrm{Re}}{C_\mathrm{M}}\frac{\partial C_\mathrm{M}}{\partial\,\mathrm{Re}}\right),\\
&\frac{\partial N}{\partial T}=-N\left(\frac{\mathrm{M}}{C_\mathrm{M}}\frac{\partial C_\mathrm{M}}{\partial \mathrm{M}}\frac{1}{2T}+\frac{\mathrm{Re}}{C_\mathrm{M}}\frac{\partial C_\mathrm{M}}{\partial\,\mathrm{Re}}\frac{1}{\mu}\frac{\partial\mu}{\partial T}\right),
\end{aligned}
\right\}\qquad(13.49)
$$

$$\left.\begin{aligned}
\frac{\partial N}{\partial w}&=\frac{\partial N}{\partial v_\theta},\\
\frac{\partial N}{\partial I}&=-\frac{N}{I}.
\end{aligned}\right\}$$

С помощью этих частных производных легко подсчитать коэффициенты a_i, b_i, c_j:

$$\left.\begin{aligned}
a_1&=\frac{\partial F}{\partial r}=\frac{\partial \Sigma}{\partial r}\sin\beta+\frac{dw}{dr}\Lambda+(v_\theta+w)\frac{\partial\Lambda}{\partial r}-v_r\frac{\partial\Delta}{\partial r}+\\
&\qquad+\left(\frac{v_\theta}{r}\pm\Omega\right)^2-2\frac{v_\theta}{r}\left(\frac{v_\theta}{r}\pm\Omega\right)+2\frac{g}{r}\left(\frac{r_0}{r}\right)^2,\\
a_2&=\frac{\partial F}{\partial\beta}=\Sigma\cos\beta+(v_\theta+w)\frac{\partial\Lambda}{\partial\beta}-v_r\frac{\partial\Lambda}{\partial\beta},\\
a_3&=\frac{\partial F}{\partial v_r}=(v_\theta+w)\frac{\partial\Lambda}{\partial v_r}-\Delta-v_r\frac{\partial\Delta}{\partial v_r},\\
a_4&=\frac{\partial F}{\partial v_\theta}=\Lambda+(v_\theta+w)\frac{\partial\Lambda}{\partial v_\theta}-v_r\frac{\partial\Delta}{\partial v_\theta}+2\left(\frac{v_\theta}{r}\pm\Omega\right),\\
a_5&=\frac{\partial F}{\partial\gamma}=(v_\theta+w)\frac{\partial\Lambda}{\partial\gamma}-v_r\frac{\partial\Delta}{\partial\gamma},\\
a_6&=\frac{\partial F}{\partial\rho}=(v_\theta+w)\frac{\partial\Lambda}{\partial\rho}-v_r\frac{\partial\Delta}{\partial\rho},\\
a_7&=\frac{\partial F}{\partial T}=(v_\theta+w)\frac{\partial\Lambda}{\partial T}-v_r\frac{\partial\Delta}{\partial T},\\
a_8&=\frac{\partial F}{\partial w}=\Lambda+(v_\theta+w)\frac{\partial\Lambda}{\partial w}-v_r\frac{\partial\Delta}{\partial w},\\
a_9&=\frac{\partial F}{\partial W}=\frac{\partial\Sigma}{\partial W}\sin\beta+(v_\theta+w)\frac{\partial\Lambda}{\partial W}-v_r\frac{\partial\Delta}{\partial W},
\end{aligned}\right\}\qquad(13.50)$$

$$\left.\begin{aligned}
b_1&=\frac{\partial G}{\partial r}=\frac{\partial\Sigma}{\partial r}\cos\beta-v_r\frac{\partial\Lambda}{\partial r}-\frac{dw}{dr}\Delta-(v_\theta+w)\frac{\partial\Delta}{\partial r}+\frac{v_rv_\theta}{r^2},\\
b_2&=\frac{\partial G}{\partial\beta}=-\Sigma\sin\beta-v_r\frac{\partial\Lambda}{\partial\beta}-(v_\theta+w)\frac{\partial\Delta}{\partial\beta},\\
b_3&=\frac{\partial G}{\partial v_r}=-\Lambda-v_r\frac{\partial\Lambda}{\partial v_r}-(v_\theta+w)\frac{\partial\Delta}{\partial v_r}-2\left(\frac{1}{2}\frac{v_\theta}{r}\pm\Omega\right),\\
b_4&=\frac{\partial G}{\partial v_\theta}=-v_r\frac{\partial\Lambda}{\partial v_\theta}-\Delta-(v_\theta+w)\frac{\partial\Delta}{\partial v_\theta}-\frac{v_r}{r},\\
b_5&=\frac{\partial G}{\partial\gamma}=-v_r\frac{\partial\Lambda}{\partial\gamma}-(v_\theta+w)\frac{\partial\Delta}{\partial\gamma},
\end{aligned}\right\}\qquad(13.51)$$

$$
\left.
\begin{aligned}
b_6 &= \frac{\partial G}{\partial \rho} = -v_r \frac{\partial \Lambda}{\partial \rho} - (v_\theta + w) \frac{\partial \Delta}{\partial \rho}, \\
b_7 &= \frac{\partial G}{\partial T} = -v_r \frac{\partial \Lambda}{\partial T} - (v_\theta + w) \frac{\partial \Delta}{\partial T}, \\
b_8 &= \frac{\partial G}{\partial w} = -v_r \frac{\partial \Lambda}{\partial w} - \Delta - (v_\theta + w) \frac{\partial \Delta}{\partial w}, \\
b_9 &= \frac{\partial G}{\partial W} = \frac{\partial \Sigma}{\partial W} \cos\beta - v_r \frac{\partial \Lambda}{\partial W} - (v_\theta + w) \frac{\partial \Delta}{\partial W},
\end{aligned}
\right\}
$$

$$
\left.
\begin{aligned}
c_1 &= \frac{\partial H}{\partial r} = -\frac{1}{r^2} \Big[\Sigma \cos\beta - v_r \Lambda - (v_\theta + w) \Delta - \\
&\qquad - 2v_r \left(\frac{v_\theta}{r} \pm \Omega \right) \Big] + \frac{1}{r} \Big[\frac{\partial \Sigma}{\partial r} \cos\beta - v_r \frac{\partial \Lambda}{\partial r} - \\
&\qquad\qquad - (v_\theta + w) \frac{\partial \Delta}{\partial r} - \frac{dw}{dr} \Delta + 2 \frac{v_r v_\theta}{r^2} \Big] + \frac{\partial N}{\partial r}, \\
c_2 &= \frac{\partial H}{\partial \beta} = \frac{1}{r} \Big[-\Sigma \sin\beta - v_r \frac{\partial \Lambda}{\partial \beta} - (v_\theta + w) \frac{\partial \Delta}{\partial \beta} \Big] + \frac{\partial N}{\partial \beta}, \\
c_3 &= \frac{\partial H}{\partial v_r} = \frac{1}{r} \Big[-\Lambda - v_r \frac{\partial \Lambda}{\partial v_r} - (v_\theta + w) \frac{\partial \Delta}{\partial v_r} - \\
&\qquad\qquad - 2 \left(\frac{v_\theta}{r} \pm \Omega \right) \Big] + \frac{\partial N}{\partial v_r}, \\
c_4 &= \frac{\partial H}{\partial v_\theta} = \frac{1}{r} \Big[-v_r \frac{\partial \Lambda}{\partial v_\theta} - \Delta - (v_\theta + w) \frac{\partial \Delta}{\partial v_\theta} - 2 \frac{v_r}{r} \Big] + \frac{\partial N}{\partial v_\theta}, \\
c_5 &= \frac{\partial H}{\partial \gamma} = \frac{1}{r} \Big[-v_r \frac{\partial \Lambda}{\partial \gamma} - (v_\theta + w) \frac{\partial \Delta}{\partial \gamma} \Big] + \frac{\partial N}{\partial \gamma}, \\
c_6 &= \frac{\partial H}{\partial \rho} = \frac{1}{r} \Big[-v_r \frac{\partial \Lambda}{\partial \rho} - (v_\theta + w) \frac{\partial \Delta}{\partial \rho} \Big] + \frac{\partial N}{\partial \rho}, \\
c_7 &= \frac{\partial H}{\partial T} = \frac{1}{r} \Big[-v_r \frac{\partial \Lambda}{\partial T} - (v_\theta + w) \frac{\partial \Delta}{\partial T} \Big] + \frac{\partial N}{\partial T}, \\
c_8 &= \frac{\partial H}{\partial w} = \frac{1}{r} \Big[-v_r \frac{\partial \Lambda}{\partial w} - \Delta - (v_\theta + w) \frac{\partial \Delta}{\partial w} \Big] + \frac{\partial N}{\partial w}, \\
c_9 &= \frac{\partial H}{\partial W} = \frac{1}{r} \Big[\frac{\partial \Sigma}{\partial W} \cos\beta - v_r \frac{\partial \Lambda}{\partial W} - (v_\theta + w) \frac{\partial \Delta}{\partial W} \Big] + \frac{\partial N}{\partial W}, \\
c_{10} &= \frac{\partial H}{\partial I} = \frac{\partial N}{\partial I}.
\end{aligned}
\right\} \qquad (13.52)
$$

После выключения двигателя тяга S исчезает. Таким образом, при $t > \bar{t}_1$ величина Σ и ее производные обращаются в нуль.

Глава XIV

СИНТЕЗ СИСТЕМ АВТОМАТИЧЕСКОГО РЕГУЛИРОВАНИЯ ПО ЗАДАННЫМ ТЕХНИЧЕСКИМ УСЛОВИЯМ

В предыдущих главах мы рассматривали вопросы разработки систем автоматического регулирования главным образом с точки зрения их исследования. Это означает, что мы задавали принципиальную схему системы и затем выясняли, насколько удовлетворительно эта система действует. В предыдущей главе мы в первый раз обратились к иному, более прямому подходу к задаче: сперва мы установили требуемое свойство системы, а затем подыскали необходимую систему управления, обеспечивающую желаемое свойство. В настоящей главе мы обобщим этот подход на случай произвольных систем регулирования, в которых желаемые критерии качества выражаются в виде некоторых интегралов от регулируемых переменных. Поведение системы, построенной в соответствии с заданным критерием, описывается уравнением весьма общего характера, обычно нелинейным.

Таким образом, система регулирования, созданная указанным методом, обычно оказывается нелинейной. Однако здесь нелинейность вводится преднамеренно для получения оптимальных свойств полной системы.

Математически принцип синтеза систем регулирования по заданным критериям можно описать следующим образом. В регулируемую систему мы вводим одну или несколько дополнительных переменных. Эти дополнительные переменные образованы искусственным путем, и поэтому законы их изменения не определяются внутренними физическими закономерностями системы регулирования. Условия для определения дополнительных переменных мы находим, налагая на всю систему требование, заключающееся в том, чтобы последняя удовлетворяла заданным свойствам. Эти условия физически осуществляются с по-

мощью вычислительного устройства, которым снабжается система. Этот принцип синтеза систем регулирования впервые предложили Боксенбом и Худ[1]). Изложение настоящей главы отчасти следует работе этих авторов.

14.1. Критерии качества систем автоматического регулирования. Если величина y служит выходом системы автоматического регулирования, то есть основания ожидать, что мера качества отработки всей системы выражается интегралом по времени от некоторой функции f переменного y. Тогда критерий качества заключается в том, чтобы этот интеграл принимал минимальное значение или имел постоянную величину, т. е.

$$\int_0^{t_1} f(y)\,dt = \text{const} \quad \text{или } \min, \tag{14.1}$$

или, в частности,

$$\int_0^{t_1} (y - y_s)^2 dt = \text{const} \quad \text{или} \quad \min\,^{2)}, \tag{14.2}$$

где t_1 — момент окончания переходного процесса, а y_s — желаемый закон изменения переменной y. Выражение (14.2) измеряет ошибку по переменной y с помощью интеграла от квадрата отклонения на протяжении времени существования этого отклонения и, таким образом, служит мерой среднеквадратичной ошибки переменной y по отношению к ее желаемому течению. Можно рассматривать и другой вид критерия качества системы, согласно которому продолжительность определенного процесса должна быть минимальной или постоянной, т. е.

$$\int_0^{t_1} dt = \text{const} \quad \text{или} \quad \min. \tag{14.3}$$

---

[1]) Boksenbom A. S., Hood R., NACA TR 1068 (1952).
[2]) См. [13, 22, 72—74]. — *Прим. перев.*

Если исходить только из одного критерия типа (14.1), то мы найдем функцию f в виде $f(y) = \text{const}$. Этот результат закономерен, поскольку функцию $f(y)$ обычно можно определить в виде постоянной, если другие переменные системы не подчинены каким-либо дополнительным критериям. Однако другие переменные системы связаны обычно некоторыми ограничениями, и условия, выражающие эти ограничения, должны быть включены в первоначальный критерий. Так, например, один из возможных критериев можно записать в следующем виде

$$\left.\begin{aligned} &\int_0^{t_1} (y - y_s)^2\, dt = \min \\ \text{при}\quad & \\ &\int_0^{t_1} f(z)\, dt = \text{const}. \end{aligned}\right\} \qquad (14.4)$$

Если, например, y—угловая скорость двигателя, а z—характеристическая температура газовой турбины двигателя, то, в силу критерия (14.4), желательно спроектировать такой регулятор, чтобы при определенном значении интеграла от функции температуры интеграл от квадрата отклонения скорости имел наименьшее значение. Этот критерий можно использовать, например, в том случае, когда известно, что превышение температуры допустимо в течение определенного промежутка времени, а на протяжении переходного процесса желательно уменьшить до минимума среднеквадратичную ошибку по скорости. Тогда интеграл от функции z выражает полное количество тепла, поступившее на лопатки турбины.

Как показывает общая теория, имеется возможность сочетать сколько угодно критериев типов (14.1)—(14.4) и построить соответствующую систему регулирования, которая автоматически удовлетворяет всем этим критериям одновременно.

Другая сторона этих критериев качества регулирования связана с определением верхних пределов интегралов в левых частях равенств (14.1)—(14.4). Здесь возникает задача о выборе промежутка времени, в котором эти интегралы

должны принимать минимальное значение или быть постоянными. Целесообразно выбрать такой промежуток времени, на протяжении которого существенные внешние возмущения сохраняют постоянные значения и в течение которого регулируемая система переходит от одного существенного режима к другому. Существенными внешними возмущениями мы называем такие возмущения, которые система регулирования не в состоянии немедленно компенсировать. Если в течение промежутка времени, соответствующего пределам интеграла, с помощью которого формулируется данный критерий, возникает существенное внешнее возмущение, то нельзя указать какой-либо физически осуществимой системы, способной предвидеть это возмущение в том смысле, чтобы сработать надлежащим образом до возникновения возмущения. Существенным режимом системы мы называем любой ее режим, характеризующийся только теми переменными, которые должны быть непрерывными. В случае турбореактивного двигателя, в котором переходный процесс может описываться дифференциальным уравнением первого порядка, режим системы определяется угловой скоростью двигателя. Если при этом следует учитывать состояние системы подачи топлива или если температура реагирует на изменения скорости не мгновенно, то для описания существенного режима двигателя необходимо задать как скорость, так и ускорение, что мы сейчас и рассмотрим.

Система регулирования, получающаяся в результате применения любого метода синтеза, должна быть физически осуществимой, причем это требование имеет две стороны. Во-первых, не исключена возможность задания критериев, или неосуществимых ни в какой системе, или несовместимых между собой. Если мы будем исходить из таких критериев, то неосуществимость системы проявится или в форме требования строить закон регулирования на основании будущих значений регулируемых переменных, или в невозможности удовлетворить граничным условиям некоторого дифференциального уравнения. В большинстве случаев ясное понимание существа используемых критериев и характера регулируемой системы позволяет предотвратить подобные противоречия.

Вторая сторона требования физической осуществимости имеет чисто математический характер. Ставится зада-

ча описать регулятор или систему регулирования (т. е. получить дифференциальное уравнение), удовлетворяющие заданному критерию качества регулирования и всем необходимым граничным условиям, которые обнаруживаются при выводе соответствующего дифференциального уравнения. Хотя математическое решение задачи может быть связано с дифференцированием или интегрированием этого дифференциального уравнения, физическое решение требует, чтобы это дифференциальное уравнение само удовлетворяло граничным условиям и, стало быть, чтобы его решение не включало неопределенных постоянных интегрирования. Таким образом, такие выражения, как

$$\dot{y} = c\dot{x}$$

и

$$y = cx,$$

не обязательно равносильны как выражения, описывающие некоторую часть системы регулирования, ибо эти выражения отличаются на неопределенную постоянную интегрирования. В линейных устойчивых системах влияние этой постоянной становится пренебрежимо малым, но в общем случае нелинейных систем, рассматриваемых здесь, эту постоянную необходимо учитывать.

14. 2. Задача обеспечения устойчивости. Требование устойчивости выражает тот специфический критерий, который не учитывается в расчете основной системы регулирования, коль скоро речь идет о качестве переходных процессов. Здесь мы встречаемся с тем же положением, которое имело место в последней главе, где требование устойчивости также не учитывалось в силу того обстоятельства, что некоторые специальные условия, наложенные на полную систему, уже обеспечивали ее удовлетворительное функционирование. Однако обычно оказывается необходимым добавлять к системе регулирования стабилизирующее устройство, которое не вступает в действие до окончания переходного процесса. Поэтому такое стабилизирующее устройство не будет влиять на способность системы отвечать остальным критериям, на основании которых ее

проектируют. Это устройство можно описать следующим образом: в системе первого порядка, если

$$\left.\begin{aligned} y &= y_s, \\ \dot{y} &= 0. \end{aligned}\right\} \qquad (14.5)$$

то

В системе второго порядка, если

$$\left.\begin{aligned} y &= y_s, \\ \dot{y} = 0 \quad &\text{и} \quad \ddot{y} = 0. \end{aligned}\right\} \qquad (14.6)$$

то

При подключении к системе регулирования этого стабилизирующего устройства система приобретает способность работать в двух различных режимах; такая система принадлежит поэтому к классу систем, работающих по нескольким законам управления (см. п. 10.9). Во время переходного процесса действует главная система регулирования, обеспечивающая заданное качество этого переходного процесса. В конце переходного процесса управление переключается на вторую систему, характеризуемую условиями (14.5) или (14.6) и предназначенную для обеспечения устойчивости системы по завершении переходного процесса, т. е. для предотвращения ухода системы от заданного ей режима[1]).

14.3. Общая теория для систем первого порядка. Исходя из критериев качества регулирования, описываемых соотношениями (14.1) – (14.4), мы можем получить уравнения, характеризующие систему регулирования, следующим образом: если в рассматриваемом перечне критериев один из интегралов должен принимать минимальное значение при условии, что другие интегралы должны быть постоянными, то, как известно из вариационного исчисления,

---

[1]) В системах, приближающихся к установившемуся режиму асимптотически, обычно под моментом окончания переходного процесса понимают момент, когда система подойдет к заданному состоянию достаточно близко, т. е. когда все непрерывные координаты системы будут близки к установившимся значениям. — *Прим. перев.*

достаточно потребовать, чтобы

$$\int_0^{t_1} f(y)\,dt + \lambda_1 \int_0^{t_1} (y-y_s)^2 dt + \lambda_2 \int_0^{t_1} f_0(z)\,dt + \lambda_3 \int_0^{t_1} dt = \min$$

или

$$\int_0^{t_1} [f(y) + \lambda_1 (y-y_s)^2 + \lambda_2 f_0(z) + \lambda_3]\,dt = \min. \qquad (14.7)$$

Величины λ_1, λ_2, λ_3 представляют собой неопределенные параметры, служащие параметрами настройки системы регулирования; их конкретные значения определяются выбором постоянных, которым должны равняться упомянутые выше интегралы. Метод λ-множителей широко используется в задачах такого типа, когда разыскиваются условия минимума при определенных связях. В самом деле, задача далеко не всегда выражается в интегральной форме; такой метод применим к любому функциональному или дифференциальному соотношению между переменными [1]).

Уравнению (14.7) можно придать весьма общий характер, если включить в него все мыслимые ограничения. Если в дальнейшем какой-либо из рассматриваемых критериев не используется, то в окончательных уравнениях соответствующий множитель λ полагают равным нулю. В тех случаях, когда значение какой-либо из постоянных, которой должен равняться один из интегралов, выражающий тот или иной критерий, равно нулю, соответствующий множитель $\lambda \to \infty$.

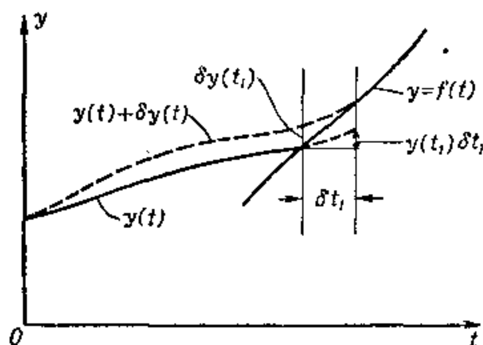
Если регулируемая система является системой первого порядка с постоянными параметрами и обладает одной существенной выходной координатой y, то между переменными y и z должна существовать связь вида $z = z(y, \dot{y})$. В этих условиях уравнение (14.7) в общем случае принимает вид

$$\int_0^{t_1} F(y, \dot{y})\,dt = \min, \qquad (14.8)$$

[1]) Lanczos C., The Variational Principles of Mechanics, Toronto, 1946.

где F — непрерывная функция переменных $\dot{y}$ и y, а y — непрерывная функция времени t. Отметим, что функция F не зависит от времени t в явном виде.

Будем рассматривать $y(t)$ как решение, т. е. пусть $y(t)$ является одной из выходных функций среди всех



Фиг. 118

допустимых выходных функций, удовлетворяющих условию (14.8). Для нахождения $y(t)$ построим решение $y(t) + \varepsilon\delta y(t)$, близкое к данному. Функция $\delta y(t)$ является произвольной, а ε — малый параметр. Если функция $y(t)$ удовлетворяет условию (14.8), то

$$\left[\frac{d}{d\varepsilon}\int_0^{t_1+\varepsilon\delta t_1} F(y+\varepsilon\delta y,\ \dot{y}+\varepsilon\delta\dot{y})\,dt\right]_{\varepsilon=0}=0$$

или

$$\int_0^{t_1}\frac{\partial F}{\partial y}\delta y\,dt+\int_0^{t_1}\frac{\partial F}{\partial\dot{y}}\delta\dot{y}\,dt+F(t_1)\,\delta t_1=0. \qquad (14.9)$$

Вариация δt_1 входит в это равенство благодаря тому, что верхний предел интеграла в левой части равенства (14.8) не фиксирован, а лежит на кривой $y=f(t)$ (фиг. 118)[1]. Это обстоятельство позволяет перемещать граничные условия задачи от значений, соответствующих

[1]) То есть рассматривается так называемая вариационная задача с подвижными концами [75]. — *Прим. перев.*

одному режиму системы, к значениям, соответствующим другому ее режиму, что уже отмечалось выше. Выполнив в левой части (14.9) интегрирование по частям, получим

$$\int_0^{t_1}\left[\frac{\partial F}{\partial y}-\frac{d}{dt}\left(\frac{\partial F}{\partial \dot y}\right)\right]\delta y\,dt+$$
$$+\left(\frac{\partial F}{\partial \dot y}\right)_{t_1}\delta y(t_1)-\frac{\partial F}{\partial \dot y}\delta y(0)+F(t_1)\delta t_1=0.$$

Соотношение между величинами $\delta y(t_1)$ и δt_1 легко найти с помощью условия на конце, т. е. с помощью условия

$$\dot y\delta t_1+\delta y(t_1)=f'(t_1)\delta t_1.$$

Далее, исключая вариацию $\delta y(t_1)$ и учитывая, что вариация δt принимает произвольные значения, приходим к равенствам

$$\int_0^{t_1}\left[\frac{\partial F}{\partial y}-\frac{d}{dt}\left(\frac{\partial F}{\partial \dot y}\right)\right]\delta y\,dt=0 \qquad (14.10)$$

и

$$\delta t_1\left\{F(t_1)+\left(\frac{\partial F}{\partial \dot y}\right)_{t_1}\left[f'(t_1)-\dot y(t_1)\right]\right\}-\left(\frac{\partial F}{\partial \dot y}\right)_0\delta y(0)=0. \qquad (14.11)$$

За промежуток времени, на протяжении которого должен выполняться критерий (14.8), мы принимаем продолжительность перехода системы от одного существенного режима к другому, что в данном случае характеризуется переходом от одного определенного значения y к другому определенному значению y. Таким образом, конечная часть кривой $y=f(t)$ должна представлять собой прямую линию при $f(t)=\text{const}$, откуда следует, что

$$\left.\begin{aligned}\delta y(0)&=0,\\ f'(t_1)&=0.\end{aligned}\right\} \qquad (14.12)$$

Следовательно, уравнения (14.10) и (14.11) принимают вид

$$\frac{\partial F}{\partial y}=\frac{d}{dt}\left(\frac{\partial F}{\partial \dot y}\right) \qquad (14.13)$$

и

$$F(t_1) = \dot{y}(t_1)\left(\frac{\partial F}{\partial \dot{y}}\right)_{t_1}, \text{ если } \left(\frac{\partial F}{\partial \dot{y}}\right)_0 \text{ — конечная величина.} \tag{14.14}$$

Так как $\delta y(0) = 0$, то уравнение (14.13) не обязательно должно удовлетворяться при $t = 0$. Единственные условия, которые должны выполняться в момент $t = 0$, заключаются в том, чтобы производная $(\partial F/\partial \dot{y})_0$ была конечной величиной, а функция y была непрерывной. В момент начала нового переходного процесса функции $\dot{y}$, F, $\partial F/\partial y$ и $\partial F/\partial \dot{y}$ должны претерпевать разрыв, в то время как в других точках $(0 < t < t_1)$ производная $\partial F/\partial \dot{y}$ будет, в силу уравнения (14.13), непрерывной.

Уравнение (14.13) представляет собой дифференциальное уравнение относительно функции $y(t)$, удовлетворяющее исходному критерию (14.8). Это так называемое дифференциальное уравнение Эйлера — Лагранжа нашей вариационной задачи.

В задаче, которую мы здесь рассматриваем, функция F не зависит явно от времени t. Поэтому мы можем сразу же найти первый интеграл этого уравнения. Первый интеграл уравнения (14.13), удовлетворяющий условию (14.14) на конце, имеет вид

$$F(y, \dot{y}) = \dot{y}\frac{\partial F}{\partial \dot{y}}. \tag{14.15}$$

Дифференцируя по t обе части этого равенства, получим

$$\frac{\partial F}{\partial y}\dot{y} + \frac{\partial F}{\partial \dot{y}}\ddot{y} = \ddot{y}\frac{\partial F}{\partial \dot{y}} + \dot{y}\frac{d}{dt}\left(\frac{\partial F}{\partial \dot{y}}\right).$$

Коль скоро функции y, $\partial F/\partial \dot{y}$ и т. д. непрерывны, из последнего соотношения следует, что

$$\dot{y}\left[\frac{\partial F}{\partial y} - \frac{d}{dt}\left(\frac{\partial F}{\partial \dot{y}}\right)\right] = 0.$$

Следовательно, или $\dot{y}=0$, или функция $y(t)$ удовлетворяет уравнению (14.13). Но во время переходного процесса $\dot{y}$, в общем случае, не исчезает[1]). Поэтому вместо двух условий (14.13) и (14.14), наложенных на функцию $y(t)$, мы получаем единственное уравнение (14.15).

Таким образом, уравнение (14.15) описывает ту физически осуществимую систему, свойства которой автоматически и одновременно удовлетворяют всем критериям, учтенным при образовании функции F, и при этом на протяжении того промежутка времени, на котором внешние возмущения постоянны, а система переходит от одного режима к другому.

В конце переходного процесса, когда переменная y устанавливается в соответствии с заданным режимом, необходимо включать стабилизирующее устройство системы. Действие этого устройства в идеальном случае описывается соотношениями (14.5).

14.4. Приложение общей теории к задаче о регулировании турбореактивного двигателя. В условиях, на которые обычно рассчитывается регулирующая аппаратура турбореактивного двигателя, регулируемой величиной служит угловая скорость N двигателя, определяющая его существенный режим, подлежащий стабилизации или изменению в соответствии с известными требованиями. В результате определяются и другие характеристики двигателя, как, например, характеристика тяги. На работу двигателя налагаются некоторые условия, ограничивающие избытки по скорости, по температуре, по давлению на выходе из компрессора и по факелу пламени. Обозначим через N_s заданную угловую скорость двигателя, через T – температуру смеси на входе турбины и через P — давление воздуха на выходе из компрессора. Тогда критерии качества работы этого двигателя можно выразить с по-

---

[1]) Не исчезает тождественно, т. е. ни на каком сплошном отрезке оси времени. — *Прим. перев.*

мощью следующих интегралов:

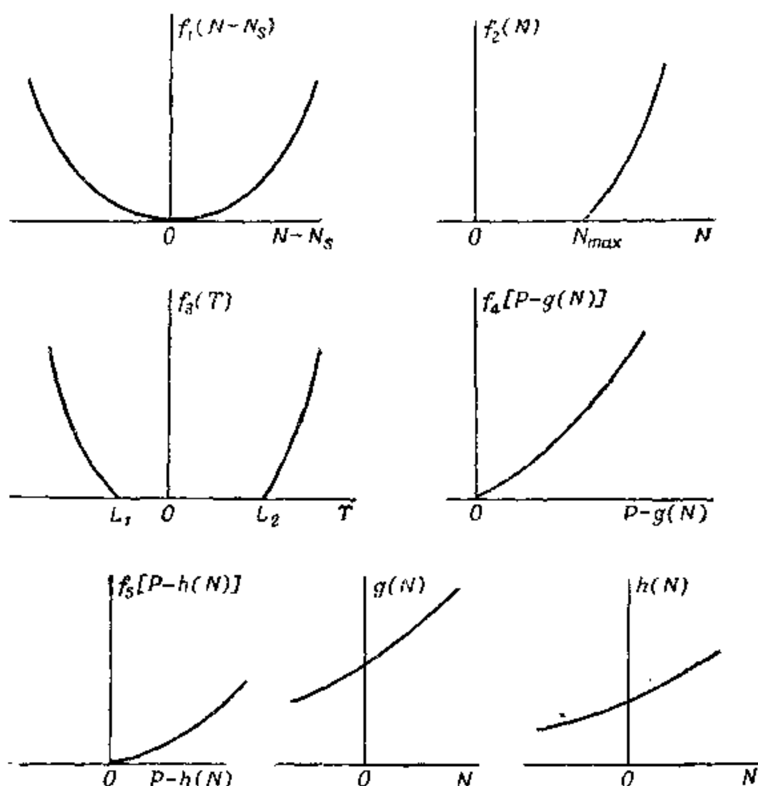
$$\left.\begin{array}{ll}
\int\limits_0^{t_1} f_1(N-N_s)\,dt & \text{критерий качества регулирования скорости,} \\
\int\limits_0^{t_1} f_2(N)\,dt & \text{критерий допустимого перерегулирования по скорости,} \\
\int\limits_0^{t_1} f_3(T)\,dt & \text{критерий допустимых отклонений температуры вверх и вниз от установившегося значения,} \\
\int\limits_0^{t_1} f_4[P-g(N)]\,dt & \text{критерий допустимых отклонений выходного давления компрессора,} \\
\int\limits_0^{t_1} f_5[P-h(N)]\,dt & \text{критерий допустимых колебаний факела пламени} \\
\text{и} & \\
\int\limits_0^{t_1} dt & \text{критерий, определяющий быстроту разгона двигателя.}
\end{array}\right\} \quad (14.16)$$

Характер этих подинтегральных функций изображен на фиг. 119. Величина $P-g(N)$ равна избытку давления на выходе из компрессора над его безопасным значением, причем $g(N)$—давление на выходе из компрессора—для каждого значения скорости меньше давления, при котором начинается помпаж, на величину, определяемую из соображений безопасности. Подобным же образом можно определить давление $h(N)$ по отношению к размерам факела. Под временем разгона понимается промежуток времени, требующийся для перевода системы из одного состояния в другое.

Пользуясь описанием турбореактивного двигателя, приведенным в п. 5.6, выпишем линеаризованные уравнения

двигателя в виде

$$\left.\begin{aligned} T &= aN + a\tau\dot{N}, \\ P &= bN + c\dot{T}, \end{aligned}\right\} \qquad (14.17)$$



Фиг. 119

где τ — постоянная времени двигателя. Подставив эти соотношения в интегралы (14.16), мы увидим, что все эти интегралы имеют вид

$$\int_0^{t_1} f(N, \dot{N})\,dt,$$

где f — непрерывная функция переменных N и $\dot{N}$, а N — непрерывная функция от t.

14.5. Регулирование скорости при ограничениях на колебания температуры. Если считать, что в нашей задаче важно только рассогласование системы по скорости, то критерий качества системы принимает вид

$$\int_0^{t_1} f(N-N_s)\,dt = \min.$$

Тогда уравнение (14.15), определяющее закон регулирования, сводится к простому уравнению

$$f_1(N-N_s) = 0.$$

Следовательно, в силу свойств функции f_1, $N=N_s$. Этот результат означает, что при отсутствии других критериев качества работы двигателя программа регулирования должна обеспечивать тождественное уничтожение ошибки по скорости, что могло бы быть физически осуществимо лишь при допущении неограниченно больших температур. Однако этот результат не соответствует выводам, сделанным ранее из уравнения (14.15), в том отношении, что N не является разрывной функцией времени. Этот пример служит лишь тривиальным частным случаем общей задачи. Но полученное решение указывает на то, что для получения физически представимой системы необходимо, чтобы критерий, подобный указанному, сопровождался некоторым дополнительным критерием.

Поэтому рассмотрим задачу регулирования скорости совместно с условием ограничения колебаний температуры, выражаемым уравнением

$$\int_0^{t_1} [f_1(N-N_s)+\lambda f_3(T)]\,dt = \min. \tag{14.18}$$

Следовательно, $F=f_1(N-N_s)+\lambda f_3(T)$. Пользуясь уравнениями (14.17), мы приведем уравнение (14.15) к виду

$$f_1(N-N_s)+\lambda f_3(T) = \lambda a\tau \dot{N} f_3'(T). \tag{14.19}$$

Таково уравнение закона регулирования во время пере-

ходного процесса. В конце этого процесса включается идеальное стабилизирующее устройство, так что, коль скоро

то
$$\left.\begin{aligned} N &= N_s, \\ \dot{N} &= 0. \end{aligned}\right\} \qquad (14.20)$$

Соотношения (14.19) и (14.20) описывают систему регулирования всю в совокупности. Имея этот результат, мы можем представить себе, что вычислительное устройство воспринимает данные от приборов, измеряющих N и T, использует запоминаемые данные λ, a и τ, умеет преобразовывать входные данные в соответствии с соотношением, связывающим расход топлива с величинами N и T, и в итоге вырабатывает сигнал, определяющий надлежащий расход топлива, в согласии с соотношением (14.19). Тогда, незадолго до того, как величина N достигнет значения N_s, включается стабилизирующее устройство, так что в конце переходного процесса будет удовлетворяться условие (14.20)[1]. В общем случае уравнение (14.19) программы регулирования нелинейно и вычислительное устройство не может быть линейным, например простой цепью типа RC.

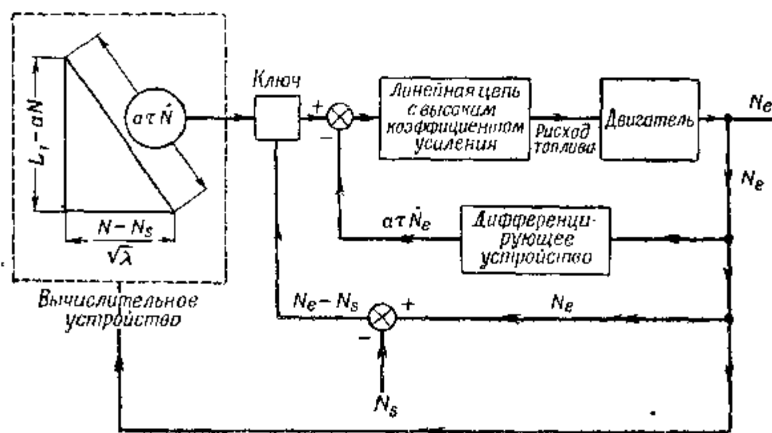
В качестве примера целесообразно задать $f_3(T) = (T - L_2)^n$ при $T > L_2$ и $f_3(T) = (L_1 - T)^n$ при $T < L_1$. В общем случае показатель n степени должен превосходить единицу, ибо при $n < 1$ величина $\dot{T}$ может обратиться в бесконечность и притом таким образом, что функция $\dot{N}$ становится разрывной и физически нереальной, даже если интеграл

$$\int_0^{t_1} f_3(T)\,dt$$

[1]) В технике широко применяется несравненно более грубый способ ограничения различных переменных параметров, основанный на использовании ограничителей (предельных регуляторов) по регулируемым величинам (скорости, давлению, температуре и т. п.): ограничители включаются каждый раз, когда данная величина достигает значения, на которое настроен предельный регулятор и которое предельно допустимо по техническим условиям. — *Прим. перев.*

ограничен. Пусть в нашем примере $n=2$ и, кроме того, $f_1(N-N_s)=(N-N_s)^2$. Поэтому мы вновь примем за меру ошибки системы среднеквадратичное отклонение регулируемой переменной от ее заданного значения. Тогда, на основании равенства (14.19),

$$\frac{(N-N_s)^2}{\lambda}+(L-aN)^2=a^2\tau^2\dot{N}^2, \tag{14.21}$$



Фиг. 120

где в случае ускоренного вращения, т. е. при $N<N_s$,

$$\dot{N}>0 \quad \text{и} \quad L=L_2. \tag{14.22}$$

В случае замедленного вращения, т. е. при $N>N_s$,

$$\dot{N}<0 \quad \text{и} \quad L=L_1. \tag{14.23}$$

Блок-схема соответствующей системы автоматического регулирования изображена на фиг. 120. Здесь N_e означает истинную угловую скорость двигателя, причем мы рассматриваем случай замедленного вращения, когда $N>N_s$. Во время переходного процесса разность N_e-N_s остается положительной, ключ между вычислительным устройством и регулятором двигателя замкнут, и действием регулятора управляет сигнал, вырабатываемый этим устройством.

Сигнал, снимаемый с вычислительного устройства, определяет величину $a\tau\dot{N}$ в соответствии с программой регулирования, выражаемой соотношением (14.21). На фиг. 120 этот сигнал схематически изображен в виде прямоугольного треугольника. Регулятор и исполнительный привод двигателя выполнены таким образом, что двигатель достаточно точно отрабатывает сигнал от вычислительного устройства. Это достигается за счет цепи с высоким коэффициентом усиления, показанной на фиг. 120. Когда же рассогласование $N_c - N_з$ по скорости падает до очень небольшой величины, вычислительное устройство отключается от регулятора. Тогда устойчивость по отношению к заданной угловой скорости $N_з$ полной системы обеспечивается стабилизирующими свойствами регулирующей цепи двигателя, и система существенно удовлетворяет условию (14.20).

Система регулирования содержит один параметр настройки — параметр λ. При любом значении λ эта система будет обеспечивать минимальное значение интеграла от квадрата ошибки по скорости, соответствующее величине интеграла от перерегулирования по температуре. Величина этого интеграла определяется значением параметра λ. Рассмотрим частный случай, для которого $aN_з = L$, т. е. разгон или замедление двигателя в процессе изменения скорости до ее значений, отвечающих границам допустимого колебания температуры, происходит согласно уравнению (14.17). Интересно отметить, что в этом частном случае, как следует из закона регулирования (14.21), $\dot{N} = 0$ при $N = N_з$. Поэтому здесь не возникает надобности в отдельном стабилизирующем устройстве, и ключ, изображенный в схеме системы автоматического регулирования на фиг. 14.3, можно удалить. В данном случае уравнение (14.21) становится линейным и записывается в виде

$$E(L - aN) = a\tau\dot{N}, \tag{14.24}$$

где

$$E = \left(1 + \frac{1}{a^2\lambda}\right)^{\frac{1}{2}}. \tag{14.25}$$

Теперь уже легко подсчитать необходимые интегралы. Например, интеграл от температуры вычисляется следующим образом:

$$\int_0^{t_1} (T-L)^2\,dt = \int_0^{t_1} (aN - L + a\tau\dot{N})^2\,dt =$$

$$= (E-1)^2 \int_0^{t_1} (L-aN)^2\,dt =$$

$$= a^2(E-1)^2 \int_0^{t_1} (N_s - N)^2\,dt = a^2(E-1)^2 \int_{N_0}^{N_s} (N_s-N)^2 \frac{dN}{\dot{N}} =$$

$$= a^3\tau(E-1)^2 \int_{N_0}^{N_s} \frac{(N_s-N)^2\,dN}{Ea(N_s-N)} = a^2\tau\frac{(E-1)^2}{E}\frac{1}{2}(N_s-N_0)^2 =$$

$$= \frac{1}{2}\frac{(E-1)^2}{E}(L-aN_0)^2.$$

Таким образом,

$$\frac{1}{\tau}\int_0^{t_1} \frac{(T-L)^2}{(L-aN_0)^2}\,dt = \frac{(E-1)^2}{2E}, \qquad (14.26)$$

где N_0 — угловая скорость двигателя в начале переходного процесса. Подобным же образом интеграл от угловой скорости имеет вид

$$\frac{a^2}{\tau}\int_0^{t_1} \left(\frac{N-N_s}{L-aN_0}\right)^2 dt = \frac{1}{2E} \qquad (14.27)$$

и, если $T_{\max}$ есть наибольшая температура,

$$\frac{T_{\max}-L}{L-aN_0} = E-1. \qquad (14.28)$$

С помощью равенства (14.24) найдем

$$Ea(N_s-N) = a\tau\frac{dN}{dt}$$

и, следовательно, характеристическое время τ^* регулируемого переходного процесса определяется отношением

$$\tau^* = \frac{\tau}{E}. \qquad (14.29)$$

Последние соотношения выражены в безразмерной форме. Наибольшая температура T_{max} достигается в начале переходного процесса.

При $E=1$ ($\lambda=\infty$) переходный процесс по температуре протекает без перерегулирования в соответствии с нашим прежним замечанием о том, что при $\lambda \to \infty$ постоянная, которой равен данный интеграл, сводится к нулю. Интеграл от угловой скорости равен 0,5, а $\tau^*=\tau$. С ростом E (т. е. с убыванием λ) величина интеграла от температуры и наибольшая температура возрастают, в то время как величина интеграла от угловой скорости и постоянная времени убывают. В качестве компромиссного значения параметра E можно выбрать $\sqrt{2}$, что соответствует $a^2\lambda=1$. Коль скоро величина E или λ выбрана, тем самым установлена конкретная программа операций, осуществляемых вычислительным устройством. С этого момента можно приступать к проектированию системы.

В общем случае, отвечающем уравнению (14.21), подсчет значений интегралов связан с несколько более громоздкими преобразованиями, но методика расчета системы регулирования остается такой же. Боксенбому и Худу удалось получить решение уравнения (14.26), т. е. найти N в функции времени t. Однако в связи с этим следует подчеркнуть, что для расчета системы регулирования не обязательно находить такое решение в явной форме. Все данные, необходимые для разработки этой системы, заложены в самом уравнении (14.24), так как это уравнение определяет конструкцию вычислительного устройства. Если это устройство системы регулирования действует в соответствии с данным уравнением, то тем самым обеспечено требуемое качество регулирования. Форма же изменения переменной N во времени, таким образом, не имеет значения. Отсюда следует, что излагаемый здесь метод проектирования заключается в «проектировании» самого нелинейного уравнения, определяющего закон изменения регулируемой величины, а не в том, чтобы вести

проектирование, исходя из решения уравнения, с помощью которого мы условились характеризовать качество системы.

14.6. Системы второго порядка с двумя степенями свободы. В случае системы второго порядка с двумя степенями свободы и с постоянными коэффициентами уравнение (14.8) принимает вид

$$\int_0^{t_1} F(y, \dot{y}, \ddot{y}, z, \dot{z}, \ddot{z})\,dt = \min, \tag{14.30}$$

где y и z суть выходные величины, являющиеся независимыми функциями времени. Уравнение (14.30) удовлетворяется при условии, что

$$\left.\begin{aligned} \frac{d}{d\varepsilon}\int_0^{t_1+\varepsilon\delta t_1} F(y+\varepsilon\delta y,\ \dot{y}+\varepsilon\delta\dot{y},\ \ddot{y}+\varepsilon\delta\ddot{y},\\ z+\varepsilon\delta z,\ \dot{z}+\varepsilon\delta\dot{z},\ \ddot{z}+\varepsilon\delta\ddot{z})\,dt = 0\\ \text{при } \varepsilon = 0. \end{aligned}\right\} \tag{14.31}$$

Промежуток времени, по которому берется интеграл (14.30), начинается со вполне определенного момента ($t = 0$), но оканчивается не в какой-либо определенный момент, а в момент времени, соответствующий кривым $y = f_1(t)$, $\dot{y} = f_2(t)$, $z = g_1(t)$ и $\dot{z} = g_2(t)$. Вариации δy и δz являются произвольными функциями времени, разумеется, независимыми одна от другой.

Выполнив преобразования, указанные в левой части равенства (14.31), получим

$$\int_0^{t_1}\left[\frac{\partial F}{\partial y}\delta y + \frac{\partial F}{\partial \dot{y}}\delta\dot{y} + \frac{\partial F}{\partial \ddot{y}}\delta\ddot{y} + \frac{\partial F}{\partial z}\delta z + \right.$$
$$\left. + \frac{\partial F}{\partial \dot{z}}\delta\dot{z} + \frac{\partial F}{\partial \ddot{z}}\delta\ddot{z}\right]dt + F(t_1)\,\delta t_1 = 0.$$

Проинтегрировав по частям, будем иметь

$$\int_0^{t_1}\left[\frac{\partial F}{\partial y} - \frac{d}{dt}\left(\frac{\partial F}{\partial \dot{y}}\right) + \frac{d^2}{dt^2}\left(\frac{\partial F}{\partial \ddot{y}}\right)\right]\delta y\,dt +$$

$$+\int_0^{t_1}\left[\frac{\partial F}{\partial z}-\frac{d}{dt}\left(\frac{\partial F}{\partial \dot z}\right)+\frac{d^2}{dt^2}\left(\frac{\partial F}{\partial \ddot z}\right)\right]\delta z\,dt+$$

$$+\left[\frac{\partial F}{\partial \dot y}\delta y+\frac{\partial F}{\partial \ddot y}\delta \dot y-\frac{d}{dt}\left(\frac{\partial F}{\partial \ddot y}\right)\delta y\right]_0^{t_1}+$$

$$+F(t_1)\,\delta t_1+\left[\frac{\partial F}{\partial \dot z}\delta z+\frac{\partial F}{\partial \ddot z}\delta \dot z-\frac{d}{dt}\left(\frac{\partial F}{\partial \ddot z}\right)\delta z\right]_0^{t_1}=0. \quad (14.32)$$

Как и выше, подинтегральные выражения и члены, соответствующие условиям на концах, должны уничтожаться по отдельности. В силу принятых условий на концах мы имеем

$$\left.\begin{aligned}\delta y(t_1)&=[f_1'(t_1)-\dot y(t_1)]\,\delta t_1,\\ \delta \dot y(t_1)&=[f_2'(t_1)-\ddot y(t_1)]\,\delta t_1,\\ \delta z(t_1)&=[g_1'(t_1)-\dot z(t_1)]\,\delta t_1,\\ \delta \dot z(t_1)&=[g_2'(t_1)-\ddot z(t_1)]\,\delta t_1.\end{aligned}\right\} \quad (14.33)$$

Три условия из равенства (14.32) представляют собой два уравнения Эйлера — Лагранжа

$$\left.\begin{aligned}\frac{\partial F}{\partial y}-\frac{d}{dt}\left(\frac{\partial F}{\partial \dot y}\right)+\frac{d^2}{dt^2}\left(\frac{\partial F}{\partial \ddot y}\right)&=0,\\ \frac{\partial F}{\partial z}-\frac{d}{dt}\left(\frac{\partial F}{\partial \dot z}\right)+\frac{d^2}{dt^2}\left(\frac{\partial F}{\partial \ddot z}\right)&=0,\end{aligned}\right\} \quad (14.34)$$

которые должны удовлетворяться одновременно, и соотношение

$$\left\{\left[F-\dot y\frac{\partial F}{\partial \dot y}+\dot y\frac{d}{dt}\left(\frac{\partial F}{\partial \ddot y}\right)-\ddot y\frac{\partial F}{\partial \ddot y}-\dot z\frac{\partial F}{\partial \dot z}+\right.\right.$$

$$\left.+\dot z\frac{d}{dt}\left(\frac{\partial F}{\partial \ddot z}\right)-\ddot z\frac{\partial F}{\partial \ddot z}\right]_{t=t_1}+f_1'(t_1)\left[\frac{\partial F}{\partial \dot y}-\frac{d}{dt}\left(\frac{\partial F}{\partial \ddot y}\right)\right]_{t=t_1}+$$

$$+f_2'(t_1)\left(\frac{\partial F}{\partial \ddot y}\right)_{t=t_1}+g_1'(t_1)\left[\frac{\partial F}{\partial \dot z}-\frac{d}{dt}\left(\frac{\partial F}{\partial \ddot z}\right)\right]_{t=t_1}+$$

$$\left.+g_2'(t_1)\left(\frac{\partial F}{\partial \ddot z}\right)_{t=t_1}\right\}\delta t_1+\delta y(0)\left[\frac{\partial F}{\partial \dot y}-\frac{d}{dt}\left(\frac{\partial F}{\partial \ddot y}\right)\right]_{t=0}+$$

$$+\delta\dot{y}(0)\left(\frac{\partial F}{\partial \ddot{y}}\right)_{t=0}+\delta z(0)\left[\frac{\partial F}{\partial \dot{z}}-\frac{d}{dt}\left(\frac{\partial F}{\partial \ddot{z}}\right)\right]_{t=0}+$$

$$+\delta\dot{z}(0)\left(\frac{\partial F}{\partial \ddot{z}}\right)_{t=0}=0. \quad (14.35)$$

Система (14.34) представляет собой систему двух уравнений Эйлера—Лагранжа, соответствующих исходному критерию (14.30). Физическое решение задачи должно удовлетворять системе (14.34) и, кроме того, условиям на концах, выражаемым соотношением (14.35). Однако, поскольку функция F не зависит от t явным образом, эти условия можно видоизменить: именно, умножив обе части первого из уравнений (14.34) на $\dot{y}$, а обе части второго на $\dot{z}$ и сложив результаты, мы получим выражение полной производной и после интегрирования придем к соотношению

$$F-\dot{y}\frac{\partial F}{\partial \dot{y}}+\dot{y}\frac{d}{dt}\left(\frac{\partial F}{\partial \ddot{y}}\right)-\ddot{y}\frac{\partial F}{\partial \ddot{y}}-$$

$$-\dot{z}\frac{\partial F}{\partial \dot{z}}+\dot{z}\frac{d}{dt}\left(\frac{\partial F}{\partial \ddot{z}}\right)-\ddot{z}\frac{\partial F}{\partial \ddot{z}}=C. \quad (14.36)$$

Так как F есть функция от $\dot{y}$ и $\ddot{y}$, то маловероятно, чтобы производные $\partial F/\partial\dot{y}$ и $\partial F/\partial\ddot{y}$ обращались в нуль или производная $\partial F/\partial\ddot{y}$ сводилась к постоянной по отношению ко времени. Поэтому разность

$$\frac{\partial F}{\partial \dot{y}}-\frac{d}{dt}\left(\frac{\partial F}{\partial \ddot{y}}\right)$$

не равна нулю. Следовательно, есть основание считать, что в соотношении (14.35)

$$\delta y(0)=\delta\dot{y}(0)=f_1'(t_1)=f_2'(t_1)=0.$$

Подобное же рассуждение применимо и к переменной z. Таким образом, есть основания ввести следующую систему

граничных условий:

$$\left.\begin{aligned} &\delta y(0)=0, && f_1'(t_1)=0,\\ &\delta\dot y(0)=0, && f_2'(t_1)=0,\\ &\delta z(0)=0, && g_1'(t_1)=0,\\ &\delta\dot z(0)=0, && g_2'(t_1)=0. \end{aligned}\right\} \qquad (14.37)$$

Эти граничные условия являются также условиями, соответствующими фиксированным начальным и конечным значениям переменных y, $\dot y$, z и $\dot z$, но при переменной продолжительности t_1 переходного процесса. Принимая во внимание условия (14.37), мы установим с помощью соотношения (14.35), что постоянная C в правой части (14.36) должна быть равна нулю. Окончательное решение задачи об обеспечении качества регулирования, соответствующего условию (14.30), определяется уравнением

$$F-\dot y\frac{\partial F}{\partial\dot y}+\dot y\frac{d}{dt}\left(\frac{\partial F}{\partial\ddot y}\right)-\ddot y\frac{\partial F}{\partial\ddot y}-\dot z\frac{\partial F}{\partial\dot z}+\dot z\frac{d}{dt}\left(\frac{\partial F}{\partial\ddot z}\right)-\ddot z\frac{\partial F}{\partial\ddot z}=0 \qquad (14.38)$$

и одним из следующих двух уравнений:

$$\left.\begin{aligned} &\frac{\partial F}{\partial y}-\frac{d}{dt}\left(\frac{\partial F}{\partial\dot y}\right)+\frac{d^2}{dt^2}\left(\frac{\partial F}{\partial\ddot y}\right)=0,\\ &\frac{\partial F}{\partial z}-\frac{d}{dt}\left(\frac{\partial F}{\partial\dot z}\right)+\frac{d^2}{dt^2}\left(\frac{\partial F}{\partial\ddot z}\right)=0. \end{aligned}\right\} \qquad (14.39)$$

Уравнения (14.38) и (14.39) образуют поэтому систему двух уравнений относительно двух неизвестных y и z. Они служат уравнениями для закона регулирования и уравнениями, определяющими действие вычислительного устройства.

Условия (14.37) на концах характеризуют исходные критерии на протяжении времени перехода системы от одного существенного режима к другому. Таким образом, коль скоро соблюдаются условия (14.37), система переходит от состояния, определяемого одним сочетанием величин y, $\dot y$, z и $\dot z$, к состоянию, определяемому другим сочетанием этих величин. Уравнение (14.38) имеет третий

порядок, а уравнение (14.39) – четвертый порядок. Поэтому, помимо четырех начальных значений переменных y, $\dot{y}$, z, $\dot{z}$, мы имеем еще возможность задать три значения функций $\dot{y}$, z, $\dot{z}$, соответствующих значению y в конце процесса, т. е. задать значения $\dot{y}=0$, $z=z_s$ и $\dot{z}=0$ при $y=y_s$. При этом все же необходимо дополнить систему стабилизирующим устройством для того, чтобы, начиная с момента окончания переходного процесса, стабилизировать значения

$$\ddot{y}=0 \qquad \text{и} \qquad \ddot{z}=0.$$

Мы видим, таким образом, что, хотя исследование рассмотренной системы носит более сложный характер по сравнению с исследованием системы первого порядка, проведенным в предыдущих разделах, к системе второго порядка все же полностью применима та же общая методика.

14.7. Задача регулирования с дополнительным условием, описываемым посредством дифференциального уравнения. Пусть y – существенная регулируемая переменная, а z — переменная, которую мы вводим в систему для обеспечения требуемого качества регулирования по переменной y[1]). Тогда динамика системы, определяемая ее физическими свойствами, доставляет одно соотношение между y и z. Это соотношение обычно выражается в виде нелинейного дифференциального уравнения, например уравнения второго порядка

$$g(y,\ \dot{y},\ \ddot{y};\ z,\ \dot{z},\ \ddot{z};\ t)=0. \tag{14.40}$$

Требования к качеству переходного процесса задаются, например, посредством условия

$$\int_0^{t_1} f(y)\,dt = \min. \tag{14.41}$$

Примером условия такого рода служит интеграл в левой части (14.2), характеризующий ошибку системы.

[1]) Например, y — угловая скорость двигателя, а z — координата заслонки, дозирующей поступление топлива в камеры горения.—*Прим. перев.*

Задача состоит в нахождении уравнения такого закон регулирования, который удовлетворял бы как уравнени (14.40), так и условию (14.41).

Математическая задача в этом случае приводитс к вариационной задаче, в которой дифференциально уравнение служит дополнительным условием. Эту задач можно также решить с помощью метода введения мнс жителя Эйлера — Лагранжа $\lambda(t)$ [1]), т. е. потребовать чтобы

$$\int_0^{t_1} F(y, \dot{y}, \ddot{y}; z, \dot{z}, \ddot{z}; t)\,dt = \min \qquad (14.42)$$

при

$$F = f(y) + \lambda(t)\, g(y, \dot{y}, \ddot{y}; z, \dot{z}, \ddot{z}; t). \qquad (14.43)$$

Единственная новая особенность задачи состоит в поя лении множителя $\lambda(t)$, зависящего от времени. Задача характеризуемая соотношением (14.42), в точности совпа дает с задачей, характеризуемой условием (14.30). По этому можно использовать все уравнения, полученные в пре дыдущем разделе. Однако теперь мы встречаемся уже с тре мя переменными: y, z и λ. Для их определения мы располагае тремя уравнениями: (14.34), (14.40) и уравнением (14.32) определяющим функцию F. Уравнение (14.40) выражае свойства, присущие самой физической системе, и поэтом ему система регулирования удовлетворяет автоматически Но чтобы удовлетворить уравнению (14.34), необходим искусственные меры. Следовательно, основа действия вы числительного устройства системы регулирования должн определяться системой этих двух уравнений. Вычислител ное устройство, при его надлежащем выполнении, воспри нимает значения существенного выхода y, перерабатывае их и затем вырабатывает непрерывный сигнал, управляю щий величиной z. Этот сигнал по z подается на вход объект регулирования и вынуждает систему вести себя в соответ ствии с условием (14.41).

---

[1]) См., например, Bolza O., Vorlesungen über Variationsrechnung гл. 11, 1909.

14. 8. Сравнение различных принципов построения систем автоматического регулирования. В предыдущей и настоящей главах мы рассматривали методы синтеза систем автоматического регулирования при задании совершенно жестких требований качества регулирования. Метод, развитый в предыдущей главе и основанный на теории возмущений, применим к линейным системам с переменными параметрами. Метод, изложенный в настоящей главе, является даже еще более общим в том отношении, что сам объект регулирования может быть и нелинейным. В таких системах общего вида эти новые методы служат единственным средством, которым мы располагаем для синтеза системы автоматического регулирования. Получаемые же в результате сложные регулирующие цепи, содержащие электромеханические вычислительные устройства, являются, повидимому, единственным логическим решением задачи. Однако методы, развитые в этих двух главах, в равной мере применимы к более простым физическим системам, рассмотренным в первых главах, т. е. к линейным системам с постоянными коэффициентами. Таким образом, в отношении этих более простых систем у нас есть две различные общие методики решения задачи об автоматическом регулировании. Сравнение двух соответствующих принципов построения систем автоматического регулирования приводит к поучительным выводам.

В гл. V мы разобрали задачу о регулировании двигателя с точки зрения обычных принципов следящих систем. В настоящей главе мы рассмотрели почти ту же задачу с помощью нового метода синтеза по заранее заданным критериям качества. Один вывод обнаруживается сразу: система регулирования, разработанная по старым методам, является линейной, и звенья регулирующей цепи могут сводиться к простым контурам типа RC, между тем как система регулирования, построенная с помощью новых методов, является нелинейной и существенным звеном в цепи регулирования служит вычислительное устройство, которое даже в его простейшей форме оказывается намного сложнее контура типа RC. Но это усложнение вводится не без выгоды: хотя система регулирования, основанная на принципах обычных следящих систем, и может быть вполне удовлетворительной по своим качествам, система регулирования, основанная на применении нелинейного вычислительного устройства, обес-

печивает оптимальное качество регулирования—никакая другая система не может лучше отвечать тем же самым техническим условиям, предъявленным к качеству системы. Вместе с тем это сравнение имеет смысл лишь при условии, что нам в точности известны требования к качеству регулирования. Например, если мы не знаем точно величины интеграла от температуры в условии (14.18), то мы вообще не в состоянии применить метод, изложенный в настоящей главе. С другой стороны, для синтеза удовлетворительной системы регулирования, если исходить из старых принципов в теории следящих систем, не требуется задания столь жестко определенных технических условий.

Разумеется, требование строгого соответствия системы режиму лучшего качества регулирования должно прийти после того, как достигнуто понимание существа оптимального качества процессов в системе. Поэтому, когда мы стремимся к созданию оптимальной системы регулирования, мы, естественно, должны располагать данными для строгого определения критериев, из которых следует исходить при синтезе системы. С этой точки зрения изложенные в последних главах новые принципы синтеза систем регулирования по заданным техническим условиям безусловно являются шагом вперед по сравнению с принципами обычной теории следящих систем; новые принципы приводят к системам регулирования повышенного качества. Следует, конечно, ожидать, что эти системы повышенного качества будут и более сложными по своей структуре.

Глава XV

ЭКСТРЕМАЛЬНОЕ РЕГУЛИРОВАНИЕ

В предшествующих главах мы рассматривали принципы синтеза систем автоматического регулирования в порядке возрастающей степени общности и трудности. Однако все изложение строилось на одном основном предположении: свойства и характеристики объекта регулирования всегда считались известными. Когда мы имеем дело с обычными линейными следящими системами, передаточные функции регулирующих устройств и других звеньев задаются до выполнения синтеза. Что касается линейных систем с переменными коэффициентами, то мы рассмотрели их на примере системы наведения ракеты дальнего действия. В этой задаче динамические и аэродинамические свойства ракеты были определены до начала синтеза регулирующих устройств. В общей задаче о синтезе систем автоматического регулирования по заданным наперед техническим условиям, разобранной в предыдущей главе, реакция системы на вариации регулирующих входных воздействий также являлась определенной. Таким образом, синтез системы регулирования основан на знании этих свойств системы. Обратная же связь служит лишь для передачи данных о состоянии выходной координаты системы на вычислительное устройство. В результате это устройство использует заложенные в нем данные о свойствах объекта регулирования для выработки «разумного» управляющего сигнала.

В настоящей главе мы хотим ослабить даже это явно элементарное требование о знании свойств объекта регулирования, для которого разрабатываются регулирующие устройства. Мы хотим ввести принцип *непрерывно ищущей и непрерывно измеряющей системы автоматического регулирования*, в которой задача синтеза не требует точного знания свойств объекта регулирования. Зато в процессе

регулирования производится измерение величин, определяющих эти свойства. В частности, мы разберем простую разновидность таких систем регулирования, называемую системой *экстремального регулирования* (регулирования на экстремум).

15. 1. Основные понятия. Независимо от возможной степени точности в работе вычислительного устройства точность отработки системы автоматического регулирования попрежнему зависит от достоверности данных, в соответствии с которыми система разработана. Если мы определим свойства объекта регулирования до выполнения всей системы автоматического регулирования, как это молчаливо предполагалось в предыдущих главах, то не сможем рассчитывать на достижение весьма высокой точности отработки системы по следующим причинам. Во-первых, в процессе изготовления системы всегда возникают небольшие отклонения от проекта. Например, крыло ракеты не может в точности совпадать с крылом опытного образца ракеты, продувавшегося в аэродинамической трубе для снятия его аэродинамических характеристик. Поэтому аэродинамические характеристики ракеты должны слегка отличаться от характеристик, определенных с помощью продувок. Во-вторых, всякая реальная система со временем претерпевает малые изменения. Это может происходить за счет естественной порчи материалов системы, вызванной их износом и усталостью или же изменением условий в среде, в которой работает система. Короче говоря, никогда нельзя *в точности* знать свойства реальных систем до наступления того момента времени, в который эти свойства нас интересуют. Поэтому, коль скоро требуется высокая степень точности отработки системы, необходимо применять принцип непрерывно ищущей системы автоматического регулирования.

Но стремление достичь требуемой точности отработки системы автоматического регулирования не является единственной причиной введенного изменения понятия регулирования; очень часто мы вынуждены прибегать к системе непрерывных поисков из-за возникновения больших непредвиденных изменений свойств системы. Мы уже ввели этот принцип при рассмотрении возмущающих влияний изме-

нений состояния атмосферы на действие системы автоматического наведения на цель ракеты дальнего действия. При этом мы использовали динамические свойства самой ракеты как средство для непрерывного измерения этих влияний. Еще более наглядным примером служит полет самолета в условиях обледенения. Образование ледяной корки на поверхности крыла и фюзеляжа приводит к изменению формы этих частей самолета. Больше того, сам характер образования таких ледяных отложений оказывается в известной мере неустановившимся и его нельзя сколько-нибудь точно предвидеть. Поэтому аэродинамические свойства самолета могут претерпевать большие изменения под влиянием обледенения, и эти изменения носят непредвиденный характер. Хуже того, все такие изменения приводят к ухудшению качества самолета, т. е. уменьшают длину пути, который самолет может пройти за счет единицы количества топлива. Таким образом, мы стремимся найти такое сочетание открытия заслонки, дозирующей подачу горючего в двигатель, числа оборотов ротора двигателя и установки триммера руля высоты, при котором самолет сможет пролететь наибольшее расстояние за счет одной единицы количества горючего, поскольку мы стремимся совершать полет в оптимальных условиях в смысле экономии дефицитного горючего. Но при решении этой задачи все известные нам ранее сведения о свойствах самолета оказываются бесполезными в связи с появлением обледенения. Следовательно, единственное решение этой задачи о разыскании экономического режима полета при указанных неблагоприятных обстоятельствах заключается в применении системы автоматического регулирования с непрерывными поисками и непрерывным измерением, т. е. системы экстремального регулирования, автоматически задающей самолету режим, отвечающий оптимальным эксплуатационным условиям, выявленным путем измерения.

Конечно, квалифицированный оператор управляет всякой машиной по принципу экстремального регулирования. Он наблюдает за показаниями приборов, регистрирующими входные и выходные величины машины, и затем на основании своих знаний и опыта принимает решение о том, в какую сторону нужно перемещать рукоятки управления. В результате изменения настройки входов образуются новые

значения выходов, и оператор должен рассматривать новые показания приборов с точки зрения того, достигнут ли уже оптимальный режим, или же произошло перерегулирование. В последнем случае потребуется новое изменение настройки. Последовательная настройка входов и представляет собой процесс поисков, а обратная связь осуществляется посредством чтения показаний приборов. Однако система ручного экстремального регулирования неизбежно реагирует медленно, и в сложных случаях оператор не может обеспечить удовлетворительного качества процесса, каким бы опытным он ни был.

Задача об экстремальном регулировании была разработана Дрэйпером, Ли и Г. Лэйнингом младшим[1]). Применение экстремального регулирования к задаче об отыскании экономического режима полета самолета рассматривалось Дж. Шуллом[2]).

15. 2. Принципы оптимального регулирования. Сердцем системы экстремального регулирования является нелинейное звено, характеризующее оптимальный режим системы. Для простоты рассуждения мы предположим, что это основное звено обладает одним входом и одним выходом. На данной стадии исследования мы будем пренебрегать влиянием всяких запаздываний в системе и будем предполагать, что выход определяется только мгновенной величиной входа. Так как существует оптимальный режим, то функция, определяющая зависимость выхода от входа, имеет точку максимума y_0, x_0, показанную на фиг. 121. В дальнейшем оказывается удобным рассматривать выход и вход в системе координат с началом в точке оптимума и принять за вход величину $x+x_0$, а за выход—величину y^*+y_0. Тогда точкой оптимума будет точка $x=y^*=0$. Таким образом, цель экстремального регулирования состоит в поисках этой оптимальной точки и в стабилизации системы в непосредственной близости

---

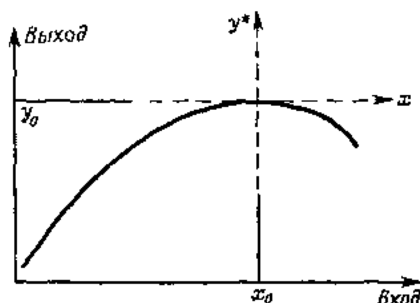
[1]) L i Y. T., Instr., 25, 72—77, 190—193, 228, 324—327, 350—352 (1952); D r a p e r C. S., L i Y. T., Principles of Optimalizing Control Systems and an Application to Internal Combustion Engine, ASME Publications (1951).

[2]) S h u l l J. R., Trans. IRE (Electronic Computers), December, 1952, 47—51.

к этой точке. В окрестности точки оптимума соотношение между x и y^* можно представить в виде зависимости

$$y^* = -kx^2. \qquad (15.1)$$

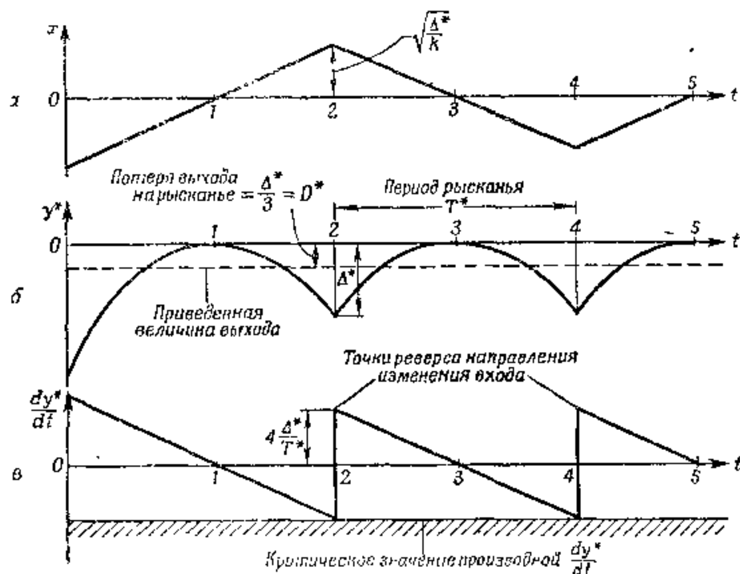
Простейший способ создания системы экстремального регулирования в принципе заключается в следующем. Пусть мы начинаем процесс регулирования при отрицательном значении входа, т. е. при значении входа, меньшем оптимального, и увеличиваем этот вход с постоянной скоростью (фиг. 122, *а*). В соответствии с этим выход будет сперва увеличиваться, затем достигнет наибольшего значения и далее будет уменьшаться (фиг. 122, *б*). Следовательно, первая производная от y^* по времени, dy^*/dt,



Фиг. 121

будет сперва положительной, затем упадет до нуля в точке *1* на фиг. 122, *в* и за этой точкой станет отрицательной. В точке *2* производная dy^*/dt достигает критической величины, при которой, в силу устройства системы регулирования, происходит реверс направления изменения входа; после этого вход x уже убывает с той же самой скоростью. В результате y^* вновь увеличивается, а величина dy^*/dt скачком изменяется от отрицательных значений к положительным. В точке *3* выход достигает максимума, а dy^*/dt вновь обращается в нуль. В точке *4* эта производная опять принимает критическое значение и вновь происходит реверс направления изменения входа. Далее этот процесс вновь повторяется и состояние системы периодически изменяется указанным образом. Об этом явлении говорят, что система «рыскает» около опти-

мального режима. Период T^* называется *периодом рысканья*. Ордината Δ^* на фиг. 122 равна наименьшему значению выхода и называется *амплитудой рысканья выхода* y^*. В силу параболического характера соотношения, графически изображенного на фиг. 121, среднее значение выхода, с учетом рысканья, на $^1/_3\,\Delta^*$ ниже его оптимального значения. Эта разность представляет собой *потерю*



Фиг. 122

D^* на рысканье, являющуюся той ценой, которую приходится платить за введение регулирования. Таким образом,

$$D^* = \frac{1}{3}\Delta^*. \qquad (15.2)$$

Другие характеристики системы можно подсчитать в функции величины Δ^* и T^*. Обращаясь к уравнению (15.1), находим, что крайние значения входа равны $\pm\sqrt{\Delta^*/k}$. Таким образом, скорость изменения входа равна $2\sqrt{\Delta^*/k}/T^*$, а критическое значение скорости изменения выхода равно $-4\Delta^*/T^*$. Поэтому, если мы зададим амплитуду Δ^* рысканья или потерю D^* на рысканье и период

T^* рысканья, то тем самым мы определим параметры всей системы. В такой системе экстремального регулирования существенны звено, задающее пробные изменения входа, звено, измеряющее и дифференцирующее выходную координату, и звено, осуществляющее переключение направления изменения входа при заранее определенных условиях. Поиски и обнаружение точки оптимума осуществляются за счет принудительных изменений входа, но из-за непрекращающихся изменений входа также возникает малая потеря D выхода. Амплитуду Δ^* рысканья желательно сделать малой, но при малых Δ^* также уменьшается и величина критического значения производной dy^*/dt, при которой происходит реверс направления изменения входа. Следовательно, такое уменьшение Δ^* приводит к увеличению опасности случайных реверсов входа, вызванных влиянием неизбежных возмущений или шума в системе. Очевидно, что коль скоро система выведена из состояния оптимального режима, то время возвращения ее в это состояние прямо пропорционально периоду T^* рысканья[1]). Поэтому желательно задавать короткий период рысканья. Но при слишком малых значениях T^* возникает трудность при отделении изменений выхода, обусловленных рысканьем системы, от изменений, вызванных случайными причинами. Мы еще вернемся к этому вопросу в следующем разделе.

Пробные изменения входа можно задавать в виде непрерывной функции времени вместо пилообразной функции, графически изображенной на фиг. 122, *а*. Например, мы можем использовать вход x, образованный путем сочетания медленно изменяющейся составляющей x_a и синусоиды постоянной амплитуды a и частоты ω, т. е.

$$x = x_a + a\sin\omega t. \tag{15.3}$$

Тогда, в силу равенства (15.1), соответствующее значение выхода y^* примет вид

$$y^* = -k\left(x_a^2 + \frac{a^2}{2}\right) - 2kx_a a\sin\omega t - \frac{ka^2}{2}\cos(2\omega t). \tag{15.4}$$

---

[1]) Точнее, это время тем больше, чем больше T^*, ибо в нелинейных системах такая зависимость, будучи монотонной, не обязательно линейна; при малых T^* эта зависимость приблизительно линейна. — *Прим. перев.*

Этот выходной сигнал можно подать на полосовой фильтр с целью подавления как медленно изменяющейся первой составляющей, так и гармоники удвоенной частоты в третьем слагаемом правой части равенства (15.4). Таким образом, через фильтр пройдет сигнал $-2kx_a \sin \omega t$. После этого профильтрованный сигнал и синусоидальный сигнал $a \sin \omega t$ подаются на множительное звено с последующим выпрямлением, с которого после умножения сигналов и выпрямления снимается величина

$$-2kx_a a^2 \sin^2 \omega t = -kx_a a^2 [1 - \cos(2\omega t)], \qquad (15.5)$$

после чего вновь отфильтровывается гармоника удвоенной частоты. Тогда, в конечном счете, мы получаем сигнал $-ka^2 x_a$, которым можно воспользоваться для изменения составляющей входа так, чтобы удовлетворялось уравнение

$$\alpha \frac{dx_a}{dt} = -ka^2 x_a. \qquad (15.6)$$

В результате переменная величина x_a стремится к нулю вместе с временем затухания $2T^*$:

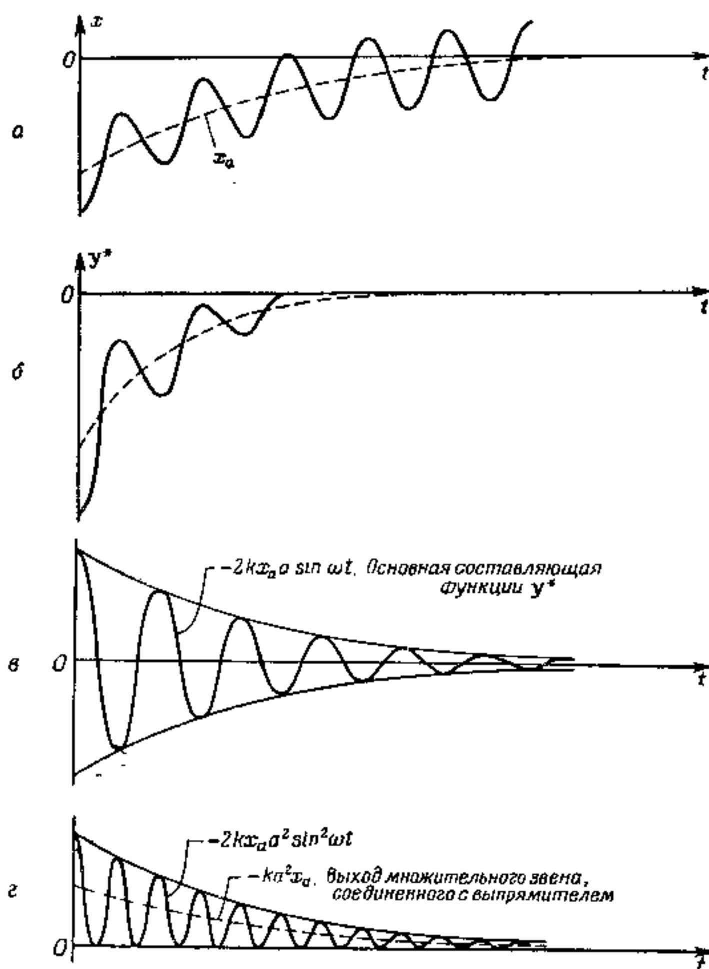
$$2T^* = \frac{\alpha}{ka^2}. \qquad (15.7)$$

В силу параболического характера соотношения, связывающего вход и выход, время затухания выхода равно T^*. Поэтому такая система регулирования будет также разыскивать точку оптимума и приближаться к ней асимптотически. Регулирующее действие при этом непрерывном пробном сигнале показано на фиг. 123. На фиг. 123, *в* изображен график выходного сигнала после прохождения через фильтр, а на фиг. 123, *г* показано влияние множительного звена с выпрямителем.

Когда система находится вблизи точки оптимума, то благодаря синусоидальному колебанию входа выход равен $-ka^2 \sin^2 \omega t$. Поэтому и в этом случае наблюдается потеря выхода $D^* = ka^2/2$. Чтобы уровень потерь был низким, амплитуда пробного колебания на входе должна быть малой. Но это уменьшение амплитуды снова оказывается ограниченным из-за соображений, связанных с влиянием случайных возмущений и шума в системе. Амплитуда Δ^*

рысканья на входе равна ka^2 и, таким образом,

$$D^* = \frac{1}{2}\Delta^*. \tag{15.8}$$



Фиг. 123

Как показывает соотношение (15.7), параметр α настройки колебания входного сигнала определяется величиной постоянной T^* времени и величиной постоянной D^* в соот-

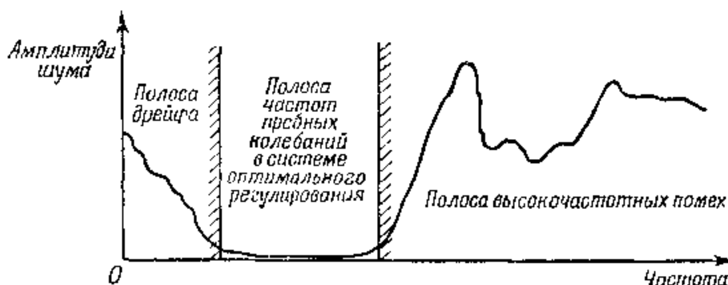
ветствии с равенством

$$\alpha = 4D^*T^* = 2\Delta^*T^*. \qquad (15.9)$$

Выпрямленная составляющая $-ka^2x_a$, определяемая соотношением (15.5) и полученная путем переработки пробного входного сигнала (15.3), служит теперь мерой отклонения входа от оптимального входа. Возмущение входа с помощью непрерывного колебания, согласно соотношению (15.6), является лишь одной из многих возможных форм использования этого сигнала. Очевидно, что с помощью этого сигнала можно также заставить переменную x_a изменяться по закону пилообразной функции путем наложения входа с постоянной скоростью изменения на синусоидальное колебание с реверсированием направления изменения входа при критических значениях сигнала $(-ka^2x_a)$. Следовательно, в таком оптимальном регулировании процесс рысканья сводится к двум раздельным колебаниям: низкочастотной составляющей в изменении x_a и высокочастотной составляющей, обусловленной синусоидальным колебанием на входе.

15. 3. **Влияние помех.** Исследование свойств систем экстремального регулирования, проведенное выше, указывает на важность уменьшения амплитуды пробных колебаний входа и постоянной T^* времени. Однако действительная разработка системы ограничивается в этом отношении влиянием шумов и помех, неизбежных в физических системах. Для замера колебаний выхода, порождаемых колебаниями входа, предназначенными для оценки степени удаления системы от точки оптимума, необходимо, чтобы эти колебания выхода во времени складывались из составляющих, принадлежащих к той области частот, в которой их заведомо можно отличить от составляющей колебаний на выходе, обусловленной воздействием шума и помех. Относительные амплитуды составляющей выхода, обусловленной воздействием помех, можно отложить на графике в функции частоты. Этот спектр помех на выходе системы обычно содержит полосу низких частот, определяемую дрейфом элементов системы, и полосу высоких частот (фиг. 124). Между этими двумя полосами обычно лежит полоса частот, относительно свободная от шумовых влияний.

Если выполнить систему экстремального регулирования с расчетом на то, чтобы частоты пробных колебаний на входе принадлежали к этой последней полосе, то представляется возможным уменьшить амплитуду этих пробных колебаний без опасения потерять эти колебания среди помех. Следовательно, в общем случае пробные функции должны складываться из составляющих, достаточно быстрых для того, чтобы не смешивать их с колебаниями дрейфа элементов системы, и в то же время достаточно медленных для того, чтобы не смешивать их с высокочастотным шумом.



Фиг. 121

Эти соображения о влиянии шума оттеняют трудности, возникающие при разработке обоих типов оптимальных систем регулирования, рассмотренных в предыдущем разделе. В системе первого вида, в которой вход изменяется по закону пилообразной функции, в качестве управляющего сигнала для реверса используется производная по времени от выхода. Если выход содержит случайные помехи, то при дифференцировании по времени влияние относительных амплитуд высокочастотных составляющих возрастает; таким образом происходит сужение полосы частот, пригодных для использования в системе оптимального регулирования. Это серьезный недостаток. Второй вид системы экстремального регулирования, в котором применяется непрерывная синусоидальная пробная функция, требует более широкой полосы частот, свободных от шумов, так как, помимо изменения низкочастотной составляющей x_a, имеется синусоидальное колебание с собственной, более высокой, частотой ω. Таким образом, если объект регули-

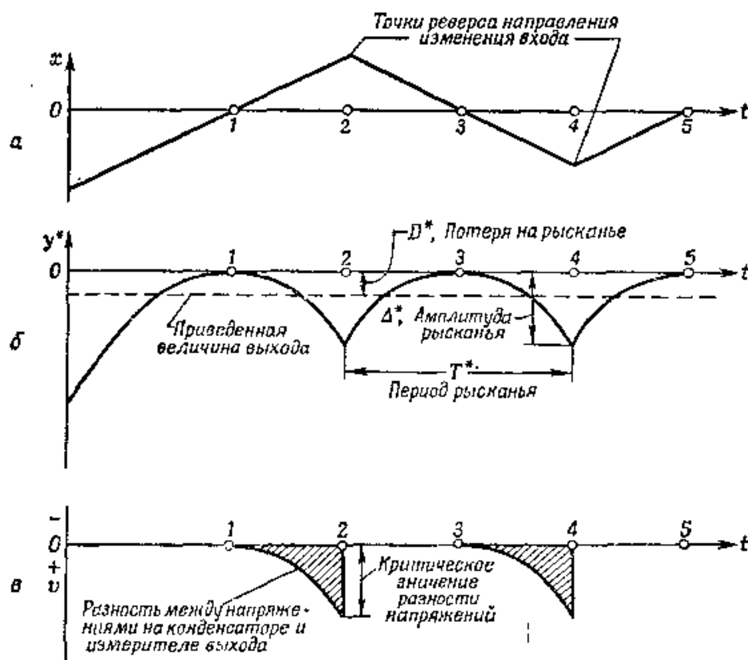
рования обладает лишь узкой полосой частот, свободных от помех, то ни одну из двух рассмотренных до сих пор систем оптимального регулирования нелегко применить. Более пригодна так называемая позиционная система экстремального регулирования, к рассмотрению которой мы и перейдем.

15. 4. Позиционная система экстремального регулирования. В позиционной системе экстремального регулирования изменение входа осуществляется по тому же закону, что и в рассмотренной выше системе оптимального регулирования первого вида, т. е. вход изменяется с постоянной скоростью при периодических переключениях направления его изменения.

Существенное улучшение состоит здесь в способе выработки сигнала реверсирования. Реверс входа должен производиться, когда выход минует его максимум и уменьшится до величины, близкой к амплитуде рысканья. Именно это обстоятельство используется для осуществления реверса входа. В главных чертах это можно выполнить следующим образом. Мерой величины выхода y^* служит некоторое электрическое напряжение. Это напряжение приложено к пластинам конденсатора через вентиль, позволяющий только заряжать конденсатор и препятствующий его разрядке. Тогда напряжение на конденсаторе следит за выходом y^* вплоть до достижения наибольшего значения. Когда же величина входа x перейдет за ее оптимальное значение, выход y^* начнет убывать. Но напряжение на конденсаторе будет оставаться равным его наибольшему значению, между напряжением на конденсаторе и напряжением выхода будет существовать разность напряжений v. Эта разность напряжений v будет расти до тех пор, пока выход не достигнет величины, равной амплитуде Δ^* рысканья. В этот момент сработает реле, которое и произведет реверс, и одновременно конденсатор разрядится до напряжения, соответствующего значению y^* в этот же момент. В итоге действие системы оптимального регулирования можно представить в виде графиков, изображенных на фиг. 125.

Амплитуда Δ^* рысканья и потеря D^* на рысканье связаны здесь тем же соотношением (15.2). Крайние значения входа попрежнему равны $\pm\sqrt{\Delta^*/k}$, а скорость изменения

входа, как и раньше, равна $2\sqrt{\Delta^*/k}/T^*$. Очевидно, что выход позиционной системы экстремального регулирования обладает только одной существенной частотой, определяемой периодом T^* рысканья, а реверс выполняется без применения дифференцирования. Таким образом, эта система



Фиг. 125

специально приспособлена для объектов регулирования, имеющих узкую полосу частот, свободную от шумов. Дальнейшего усовершенствования систем экстремального регулирования в этом направлении можно достичь, производя переключение входа не непосредственно от v, разности между напряжениями на конденсаторе и на измерителе выхода, а по интегралу от v по времени. В этом случае высокочастотные помехи будут подавлены и появится возможность дальнейшего уменьшения амплитуды рысканья и потерь на рысканье без случайных реверсов входа.

15. 5. Влияние динамических явлений. Рассуждения, проведенные в предыдущих пунктах, велись в предположении, что вход и выход системы связаны соотношением (15.1), не зависящим от скорости изменения входа или от высших производных от входа. Это предположение справедливо, если реакция выхода на вход является мгновенной, т. е. происходит без малейшего запаздывания. Но это невозможно ни в одной физической системе, ибо всегда существуют инерционные и другие динамические явления. Поэтому мы должны рассматривать выход y^*, определяемый соотношением (15.1), как фиктивный «потенциальный выход», а не действительный выход, измеряемый соответствующим прибором. Величина y^* равна y только в тех случаях, когда постоянная T^* времени цепи стабилизации оптимального режима неограниченно возрастает. Соотношение между переменными y^* и y определяется динамическими явлениями. Но мы уже убеждались в том, что такие динамические явления можно достаточно точно аппроксимировать с помощью линейной системы.

При применении экстремального регулирования к двигателю внутреннего сгорания, что сделали Дрэйпер и Ли, потенциальный выход существенно совпадает с индикаторным средним давлением двигателя, тогда как истинным выходом служит действительное среднее давление двигателя. Динамические явления здесь обусловлены главным образом инерцией поршня, кривошипа и других подвижных частей двигателя. При малых изменениях режима двигателя эти динамические явления можно описать с помощью линейного дифференциального уравнения с постоянными коэффициентами. Так как мы условились отсчитывать величины входа и выхода от точки оптимума, где вход равен x_0, а выход — y_0, то физическая потенциальная величина выхода равна y^*+y_0, а физическая индикаторная величина его равна $y+y_0$. Таким образом, соотношение между физическим потенциальным выходом и физическим индикаторным выходом можно записать в виде операторного уравнения

$$y+y_0=F_0\left(\frac{d}{dt}\right)(y^*+y_0),$$

где функция F_0 обычно выражается отношением двух многочленов по степеням оператора d/dt. С точки зрения

преобразования Лапласа $F_o(s)$ является передаточной функцией.

Линейную систему, преобразующую потенциальный выход в индикаторный выход, используемый для управления изменением входа, условимся называть *выходной линейной частью* цепи экстремального регулирования. Тогда $F_o(s)$ представляет собой передаточную функцию выходной линейной части. Однако в тех случаях, когда влиянием динамических явлений можно пренебречь, т. е. когда $s=0$, потенциальный выход равен индикаторному выходу. Поэтому мы получаем условие

$$F_o(0)=1. \tag{15.10}$$

Следовательно, так как y_o — величина постоянная, мы можем упростить вид операторного уравнения, связывающего потенциальный выход с индикаторным выходом, т. е. представить его в виде

$$y=F_o\left(\frac{d}{dt}\right)y^*. \tag{15.11}$$

Подобным же образом мы можем ввести и «потенциальный вход» x^*, который представляет собой вынуждающую функцию, вырабатываемую системой экстремального регулирования, но не является истинным входом x. Соотношение между x и x^* определяется инерционными и динамическими свойствами системы управления входом. Эту систему управления входом мы будем называть *входной линейной частью* цепи оптимального регулирования. Потенциальный вход x^* и истинный вход x связаны операторным уравнением

$$x=F_i\left(\frac{d}{dt}\right)x^*. \tag{15.12}$$

Таким образом, $F_i(s)$ представляет собой передаточную функцию входной линейной части. $F_i(s)$ связана условием

$$F_i(0)=1, \tag{15.13}$$

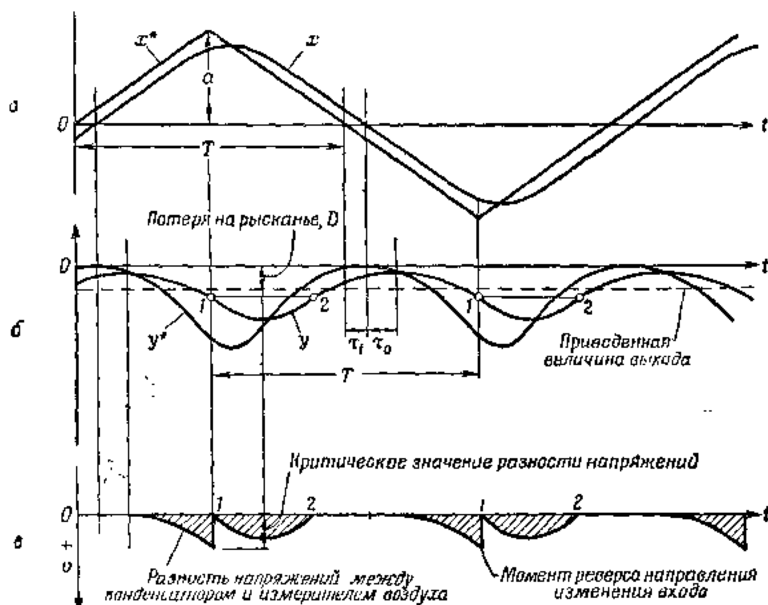
подобным условию (15.10). В результате мы можем начертить блок-схему полной системы экстремального регулирования (фиг. 126). Нелинейными звеньями этой системы служат задатчик входа, посредством которого разыскивается точка оптимума, и сам объект регулирования.

Следовательно, общее соотношение между входом x и выходом y определяется системой уравнений (15.1), (15.11)



Фиг. 126

и (15.12) и характеристикой выбранного задатчика входных колебаний. Например, если задатчик входа выполнен



Фиг. 127

по схеме позиционного устройства, рассмотренного в предыдущем разделе, то потенциальный вход x^* изображается пилообразной кривой с периодом $2T$ и амплитудой a

(фиг. 127, *a*). Введем обозначение

$$\omega_0 = \frac{2\pi}{T}. \tag{15.14}$$

Тогда x^* можно представить с помощью разложения в ряд Фурье

$$x^* = \frac{8a}{\pi^2} \sum_{n=0}^{\infty} (-1)^n \frac{1}{(2n+1)^2} \sin\left[(2n+1)\frac{\omega_0 t}{2}\right] =$$

$$= \frac{8a}{\pi^2} \sum_{n=0}^{\infty} (-1)^n \frac{1}{(2n+1)^2} \frac{1}{2i} \left[e^{\frac{2n+1}{2} i\omega_0 t} - e^{-\frac{2n+1}{2} i\omega_0 t}\right]. \tag{15.15}$$

В силу общего соотношения (2.16) истинный вход x, определяемый уравнением (15.12), можно подсчитать по формуле

$$x = \frac{8a}{\pi^2} \sum_{n=0}^{\infty} \frac{(-1)^n}{2i(2n+1)^2} \left[F_i\left(\frac{2n+1}{2} i\omega_0\right) e^{\frac{2n+1}{2} i\omega_0 t} - F_i\left(-\frac{2n+1}{2} i\omega_0\right) e^{-\frac{2n+1}{2} i\omega_0 t}\right]. \tag{15.16}$$

Потенциальный выход y^* определяется уравнением (15.1). С помощью соотношения (15.16) получим

$$y^* = \frac{16a^2k}{\pi^4} \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^{n+m}}{(2n+1)^2(2m+1)^2} \times$$
$$\times \left[F_i\left(\frac{2n+1}{2} i\omega_0\right) F_i\left(\frac{2m+1}{2} i\omega_0\right) e^{(n+m+1) i\omega_0 t} - F_i\left(\frac{2n+1}{2} i\omega_0\right) F_i\left(-\frac{2m+1}{2} i\omega_0\right) e^{(n-m) i\omega_0 t} - F_i\left(-\frac{2n+1}{2} i\omega_0\right) F_i\left(\frac{2m+1}{2} i\omega_0\right) e^{-(n-m) i\omega_0 t} + F_i\left(-\frac{2n+1}{2} i\omega_0\right) F_i\left(-\frac{2m+1}{2} i\omega_0\right) e^{-(n+m+1) i\omega_0 t}\right]. \tag{15.17}$$

Вновь применяя соотношение (2.16), найдем, наконец индикаторный выход y, определяемый уравнением (15.11) в виде

$$y=\frac{16a^2k}{\pi^4}\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\frac{(-1)^{n+m}}{(2n+1)^2(2m+1)^2}\times$$
$$\times\Big\{F_o[(n+m+1)i\omega_0]F_i\Big(\frac{2n+1}{2}i\omega_0\Big)F_i\Big(\frac{2m+1}{2}i\omega_0\Big)e^{(n+m+1)i\omega_0 t}-$$
$$-F_o[(n-m)i\omega_0]F_i\Big(\frac{2n+1}{2}i\omega_0\Big)F_i\Big(-\frac{2m+1}{2}i\omega_0\Big)e^{(n-m)i\omega_0 t}-$$
$$-F_o[-(n-m)i\omega_0]F_i\Big(-\frac{2n+1}{2}i\omega_0\Big)F_i\Big(\frac{2m+1}{2}i\omega_0\Big)\times$$
$$\times e^{-(n-m)i\omega_0 t}+F_o[-(n+m+1)i\omega_0]F_i\Big(-\frac{2n+1}{2}i\omega_0\Big)\times$$
$$\times F_i\Big(-\frac{2m+1}{2}i\omega_0\Big)e^{-(n+m+1)i\omega_0 t}\Big\}. \quad (15.18$$

Соотношения (15.17) и (15.18) ясно показывают, чт период T рысканья выхода равен лишь половине период колебания входа. Конечно, этого и следовало ожидать если иметь в виду параболический характер основног соотношения между входом и выходом.

Потерю D на рысканье мы найдем с помощью средне го по времени значения переменной y, учитывая, что здес выход отнесен к его оптимальному значению y_0. Как сле дует из соотношения (15.18), среднее значение равно сум ме второго и третьего слагаемых в правой части этог соотношения при $m=n$. Таким образом, принимая во вни мание равенство (15.10), имеем

$$D=\frac{32a^2k}{\pi^4}\sum_{n=0}^{\infty}\frac{1}{(2n+1)^4}\Big|F_i\Big(\frac{2n+1}{2}i\omega_0\Big)\Big|^2. \quad (15.19$$

Это соотношение легко проиллюстрировать на частном примере, заметив, что в отсутствии динамических явлений $F_i\equiv 1$. В этом случае сумма ряда находится легко и $D=D^*=a^2k/3=\Delta^*/3$, что совпадает с (15.2). Соотноше ние (15.19) показывает также, что среднее значение вы хода и величина потери на рысканье независимы от вы ходной линейной части. Это, конечно, следовало пред

видеть, так как установившаяся величина выхода определяется входом x и не зависит от динамики выходной линейной системы. Динамика выходной линейной части сказывается лишь в виде мелких колебаний выхода. В случае двигателя внутреннего сгорания установившимся значением выхода служит мощность двигателя. Динамика выходной линейной части определяется инерцией подвижных масс. Мощность двигателя, естественно, не зависит от инерции его подвижных масс.

В общем случае передаточных функций входной и выходной частей числовой подсчет выхода y с помощью соотношения (15.18) связан с довольно большими затруднениями. Однако в любой прикладной задаче экстремального регулирования период T рысканья обычно принимается довольно большим, с тем чтобы избежать высокочастотных помех. Тогда динамические явления не будут сказываться в сильной мере, хотя они и не будут пренебрежимо малыми. Другими словами, мы можем допустить, что отношения постоянных времени входной и выходной линейных частей к периоду рысканья малы, и вести дальнейшее исследование в соответствии с этим допущением. Например, если входная линейная часть аппроксимируется системой первого порядка с постоянной времени τ_i, т. е. если

$$F_i(s) = \frac{1}{1+\tau_i s}, \tag{15.20}$$

причем τ_i мало по сравнению с T, то безразмерная величина $\tau_i\omega_0$ также мала. При этом условии несколько первых соответственных гармоник в разложениях (15.15) и (15.16) будут попарно иметь практически одинаковые амплитуды, что отвечает соотношению (3.14). Единственное различие между этими соответственными низшими гармониками в разложениях переменных x^* и x заключается в появлении фазового сдвига, определяемого величиной τ_i. Поэтому при удаленных от точек реверса значениях x^* и x, в области которых кривизна кривых $x^*(t)$ и $x(t)$ мала, а сами значения x^* и x определяются главным образом несколькими первыми гармониками, кривая $x(t)$ отличается от кривой $x^*(t)$ лишь отставанием по оси времени на величину τ_i и не отличается от нее по

величине. При переходе от x^* к x наблюдается сглаживание угловых точек, но общая форма кривых не меняется, что и видно на фиг. 127, *а*. Если аппроксимировать системой первого порядка с постоянной времени τ_0 также и выходную линейную часть, то подобным же образом мы установим, что при переходе от y^* к y общая форма кривых не меняется, но кривая y будет отставать от кривой y^* по оси времени на величину τ_0, что и показано на фиг. 127, *б*.

Подсчитаем потерю на рысканье при входной функции, изображение которой определяется равенством (15.20). Тогда, в силу равенства (15.19),

$$D = \frac{32a^2k}{\pi^4} \sum_{n=0}^{\infty} \frac{1}{(2n+1)^4} \frac{1}{1+(2n+1)^2\left(\frac{\tau_i\omega_0}{2}\right)^2} =$$

$$= \frac{32a^2k}{\pi^4}\left[\sum_{n=0}^{\infty}\frac{1}{(2n+1)^4} - \left(\frac{\tau_i\omega_0}{2}\right)^2 \sum_{n=0}^{\infty}\frac{1}{(2n+1)^2} + \right.$$

$$\left. + \left(\frac{\tau_i\omega_0}{2}\right)^4 \sum_{n=0}^{\infty} \frac{1}{1+(2n+1)^2\left(\frac{\tau_i\omega_0}{2}\right)^2}\right];$$

но

$$\sum_{n=0}^{\infty}\frac{1}{(2n+1)^4} = \frac{\pi^4}{96}; \quad \sum_{n=0}^{\infty}\frac{1}{(2n+1)^2} = \frac{\pi^2}{8}.$$

Учитывая хорошо известное разложение гиперболического котангенса, приведенное на стр. 146, получим

$$\sum_{n=0}^{\infty} \frac{1}{1+(2n+1)^2\left(\frac{\tau_i\omega_0}{2}\right)^2} = \frac{\pi}{\tau_i\omega_0}\left(\operatorname{cth}\frac{2\pi}{\tau_i\omega_0} - \frac{1}{2}\operatorname{cth}\frac{\pi}{\tau_i\omega_0}\right).$$

В результате, используя выражение (15,14) для периода T, придем к окончательному равенству

$$D = \frac{a^2k}{3}\left[1 - 12\left(\frac{\tau_i}{T}\right)^2 + 48\left(\frac{\tau_i}{T}\right)^3\left(\operatorname{cth}\frac{T}{\tau_i} - \frac{1}{2}\operatorname{cth}\frac{T}{2\tau_i}\right)\right]. \tag{15.21}$$

Если постоянная τ_i времени значительно меньше периода T, то гиперболические котангенсы приблизительно равны единице и, следовательно,

$$D \approx \frac{a^2k}{3}\left[1-12\left(\frac{\tau_i}{T}\right)^2+24\left(\frac{\tau_i}{T}\right)^3\right]\frac{\tau_i}{T} \ll 1. \qquad (15.22)$$

Амплитуду a входа можно выразить через скорость пробного изменения входа и период T, и поэтому в позиционной системе оптимального регулирования, при входной линейной части первого порядка с постоянной τ_i времени, соотношения (15.21) и (15.22) определяют потерю D на рысканье в функции скорости пробного изменения входа и периода T рысканья. На основании этих соотношений можно было бы предположить, что потеря на рысканье убывает с ростом постоянной времени входной линейной части. Однако такое предположение обманчиво: при некотором определенном, зависящем от соображений, связанных с шумом и помехами, значении критической разности v напряжений, от которой происходит реверс входа, период T рысканья и, вследствие этого, амплитуда a окажутся бо́льшими при ненулевых значениях τ_i и τ_o, чем при нулевых их значениях. В итоге произойдет не уменьшение, а увеличение потери на рысканье.

15. 6. Расчет устойчивости режима. В любой системе автоматического регулирования свойство устойчивости означает, что система будет сохранять свое расчетное качество даже при наличии внутренних и внешних возмущений. В предыдущих главах мы рассмотрели способы обеспечения устойчивости в обычных следящих системах и других более общих системах автоматического регулирования. В системах экстремального регулирования существенная часть процесса регулирования заключается в надлежащей координации пробных колебаний на входе с отработкой системы на выходе для того, чтобы удерживать выход в достаточно близкой окрестности от точки оптимума. На этот процесс поисков точки оптимума не должны влиять ни внутренние, ни внешние возмущения. Если это свойство обеспечено благодаря удачному проектированию системы, система работает устойчиво.

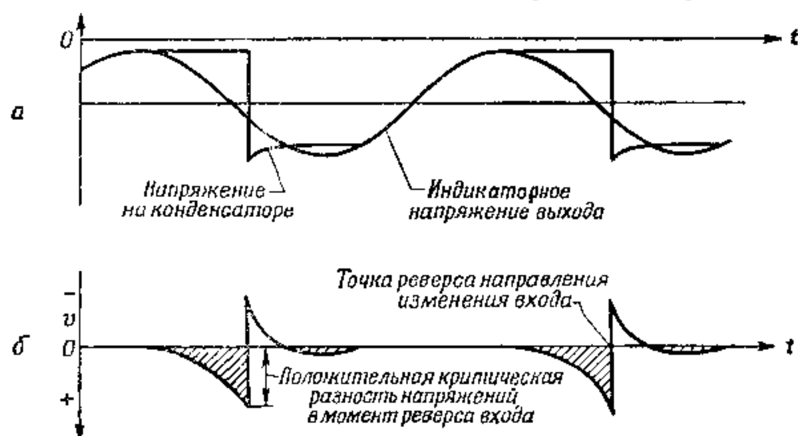
Как было описано выше, в системе экстремального регулирования с «запоминанием» максимума входной сигнал вырабатывается в процессе зарядки и разрядки конденсатора от напряжения, характеризующего величину выхода. Реверс входа происходит каждый раз, когда разность v между индицируемым напряжением выхода и напряжением на конденсаторе достигает за счет убывания выхода критической величины.

В момент реверса входа напряжение на конденсаторе за счет его разрядки падает до значения выхода, индикаторного в тот же момент. При наличии динамических явлений влияние постоянных времени входной и выходной линейных частей приведет к тому, что выход будет продолжать убывать даже после подачи сигнала реверса входа. Тогда напряжение v будет вновь возрастать и упадет лишь после того, как выход достигнет величины, соответствующей моменту подачи сигнала реверса входа. Этот процесс изображен на фиг. 127,*в*. Возникновение такого ложного положительного напряжения v в промежутке между моментами *1* и *2* (фиг. 127) нежелательно из-за угрозы ложного срабатывания задатчика пробных колебаний входа в этом промежутке времени.

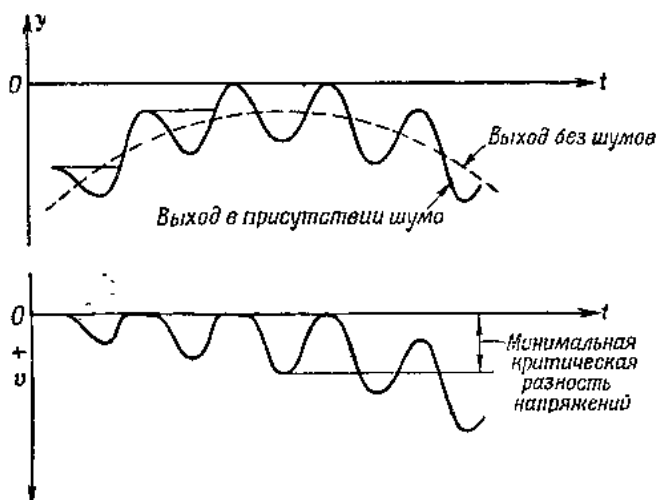
Чтобы уменьшить в значительной мере это положительное паразитное напряжение, разрядка конденсатора, выполняемая в момент реверса входа, производится до напряжения, более низкого, чем мгновенное индикаторное напряжение. Таким образом, на протяжении некоторого промежутка времени после разрядки конденсатор заряжается от индикаторного напряжения выхода. Емкость конденсатора и сопротивление электрической цепи подбираются таким образом, чтобы напряжение на конденсаторе начало приближаться к индикаторному напряжению выхода, когда выход опять начинает возрастать. Кривые изменения напряжений показаны на фиг. 128. Тем самым достигается значительное уменьшение опасного ложного положительного напряжения (фиг. 128, *б*) и улучшается устойчивость системы.

Мы уже отмечали, что возможности уменьшения амплитуды рысканья и потери на рысканье ограничены влиянием помех и шума в системе. Здесь мы вновь сталкиваемся с задачей об устойчивости работы системы: мы стремимся

избежать ложных сигналов реверса входа, появляющихся из-за действия помех. Такие ложные сигналы будут возникать при слишком малых значениях критической разности



Фиг. 128



Фиг. 129

напряжений, от которой осуществляется реверс. Это показано графически на фиг. 129, где на выход y наложен

высокочастотный синусоидальный шум. Легко видеть, что если критическая разность напряжений слишком мала, шум породит ложные сигналы реверса входа. Чтобы система работала устойчиво, критическая разность напряжений должна превосходить амплитуду помех. Следовательно, в системе экстремального регулирования потеря на рысканье не может быть меньше величины, обусловливаемой влиянием помех или шума в системе. Разумеется, если бы помеха представляла собой в реальных условиях чисто синусоидальное высокочастотное колебание постоянной амплитуды, как это показано на фиг. 129, то ее можно было бы подавить с помощью фильтра и затем работать с гораздо меньшей амплитудой рысканья. Если в действительных условиях помеха или шум имеют какую-нибудь определенную форму, то мы можем разработать надлежащий фильтр для уменьшения ограничений на амплитуду рысканья, налагаемых влиянием этих помех[1]).

---

[1]) К этим требованиям устойчивости следует добавить очевидное требование о том, чтобы автоколебания, с помощью которых система находит оптимальный режим, были устойчивыми.—*Прим. перев.*

Глава XVI

ФИЛЬТРАЦИЯ ШУМА

Во всех предыдущих рассуждениях, за исключением последней главы, мы молчаливо предполагали, что система автоматического регулирования не порождает шума и помех и что, следовательно, степень точности системы теоретически не ограничена. Как мы показали в предыдущей главе, в действительности шум и помехи искажают выходной сигнал, используемый в системе экстремального регулирования, и обусловливают основные ограничения в таких системах регулирования. Но шум и помехи присутствуют в любой технической системе, поскольку даже в «идеальных» системах имеются термодинамические флуктуации. Их влиянием на действие системы автоматического регулирования можно пренебрегать только в том случае, когда полезный сигнал является достаточно мощным по сравнению с помехами. В системах экстремального регулирования для уменьшения потерь выхода на рысканье желательно пользоваться «слабыми» сигналами и, потому задача о влиянии шума приобретает первостепенное значение. Поэтому в общем случае, коль скоро величина управляющего сигнала мала по сравнению с величиной помех, влиянием шума и помех пренеберегать нельзя.

Возмущающее влияние шума на действие системы регулирования можно свести к минимуму путем введения в систему надлежащего устройства, которое в возможно большей мере подавляет шум без уменьшения мощности полезного сигнала. Именно этот вопрос о фильтрации шума и составляет предмет рассмотрения настоящей главы.

Мы начнем с изложения теории оптимальных линейных фильтров, развитой Н. Винером[1]) и А. Н. Колмого-

[1]) N. Wiener, The Extrapo'ation, Interpolation and Smoothing of Stationary Time Series with Engincering Applications, New York, 1949.

ровым[1]). В заключительных разделах настоящей главы мы рассмотрим различные приложения и обобщения этой очень сильной теории. В рассуждениях, к которым мы переходим, весьма полезны понятия и математический аппарат теории случайных функций, развитые в главе IX.

16.1. Среднеквадратичная ошибка. Пусть $f(t)$ — управляющий сигнал, а $n(t)$ — шум. Таким образом, на вход фильтра поступает величина

$$x(t) = f(t) + n(t). \tag{16.1}$$

С выхода фильтра снимается величина $y(t)$, вырабатываемая по схеме, показанной на фиг. 130. Если фильтр



Фиг. 130

является линейным и дифференциальное уравнение, связывающее выход со входом, есть линейное уравнение с постоянными коэффициентами, то свойства фильтра полностью определяются его передаточной функцией $F(s)$. Коль скоро передаточная функция $F(s)$ известна, то реакция $h(t)$ фильтра на единичный импульс определяется соотношением (2.18). Если функция $F(s)$ имеет полюсы только в левой полуплоскости комплексной переменной s, то мы можем написать

$$h(t) = \frac{1}{2\pi i} \int\limits_{-i\infty}^{i\infty} e^{st} F(s)\, ds = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} e^{i\omega t} F(i\omega)\, d\omega. \tag{16.2}$$

Тогда выход $y(t)$ от входного воздействия $x(t)$, опреде-

[1]) Колмогоров А. Н., Изв. АН СССР, сер. матем., 5, 3—14 (1941).

ляемого равенством (16.1), равен

$$y(t)=\int_{-\infty}^{t} x(\eta)\,h(t-\eta)\,d\eta$$

в предположении, что действие выхода начинается с весьма далеких моментов времени. Положив $t-\eta=\tau$, преобразуем последний интеграл к виду

$$y(t)=\int_{0}^{\infty} x(t-\tau)\,h(\tau)\,d\tau. \tag{16.3}$$

Обозначим через $z(t)$ желаемую величину выхода, которая определяется сигналом $f(t)$ и желаемой реакцией $h_1(t)$ на единичный импульс:

$$z(t)=\int_{0}^{\infty} f(t-\tau)\,h_1(\tau)\,d\tau. \tag{16.4}$$

Функцию $h_1(t)$ можно подсчитать с помощью желаемой передаточной функции $F_1(s)$; мы получим соотношение

$$h_1(t)=\frac{1}{2\pi i}\int_{-i\infty}^{i\infty} e^{st}F_1(s)\,ds=\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega t}F_1(i\omega)\,d\omega, \tag{16.5}$$

аналогичное соотношению (16.2). Поскольку истинным выходом служит функция $y(t)$, а не $z(t)$, ошибка $e(t)$ определяется их разностью и, в силу (16.3) и (16.4),

$$e(t)=y(t)-z(t)=$$

$$=\int_{0}^{\infty}\{[f(t-\tau)+n(t-\tau)]\,h(\tau)-f(t-\tau)\,h_1(\tau)\}\,d\tau. \tag{16.6}$$

Следовательно, квадрат ошибки равен

$$e^2(t)=\int_{0}^{\infty}\int_{0}^{\infty}\{[f(t-\tau)+n(t-\tau)]\,h(\tau)-f(t-\tau)\,h_1(\tau)\}\times$$

$$\times\{[f(t-\tau')+n(t-\tau')\,h(\tau')-f(t-\tau')\,h_1(\tau')\}\,d\tau\,d\tau'. \tag{16.7}$$

Теперь мы сделаем весьма важное допущение. Так как шум $n(t)$ является случайной функцией, его можно охарактеризовать только посредством статистических свойств. Кроме того, в реальных условиях мы заранее не знаем, каким будет сигнал $f(t)$, а имеем о нем лишь общее представление. Поэтому даже в отношении сигнала мы можем располагать только данными статистического типа. Тогда мы можем оценивать статистическую ошибку с помощью среднего по множеству величины $e^2(t)$. В общем случае это среднее значение зависит от времени t. Но если мы допустим, что случайные функции $f(t)$ и $n(t)$ являются стационарными в смысле, определенном в п. 9.1, то $\overline{e^2}$ от времени не зависит. Далее, мы предположим, что среднее значение как сигнала $f(t)$, так и шума $n(t)$ равно нулю. Тогда из соотношения (16.6) со всей очевидностью вытекает, что уничтожается также и среднее значение ошибки $e(t)$. Теперь мы можем ввести следующие функции корреляции между сигналом и шумом, в точности аналогичные функциям корреляции, введенным в п. 9.2:

$$\left.\begin{aligned} \overline{f(t-\tau)\,f(t-\tau')} &= R_{ff}(\tau-\tau'), \\ \overline{f(t-\tau)\,n(t-\tau')} &= R_{fn}(\tau-\tau'), \\ \overline{n(t-\tau)\,f(t-\tau')} &= R_{nf}(\tau-\tau'), \\ \overline{n(t-\tau)\,n(t-\tau')} &= R_{nn}(\tau-\tau'). \end{aligned}\right\} \tag{16.8}$$

Эти соотношения характеризуют общий случай неисчезающей корреляции между сигналом и шумом. Очень часто функции R_{nf} и R_{fn} взаимной корреляции равны нулю и остаются только функции R_{ff} и R_{nn} автокорреляции. Функции автокорреляции представляют собой симметрические функции аргумента. Функции взаимной корреляции не являются симметрическими, а в силу их определения удовлетворяют следующим соотношениям:

$$\left.\begin{aligned} R_{fn}(\tau'-\tau) &= R_{nf}(\tau-\tau'), \\ R_{nf}(\tau'-\tau) &= R_{fn}(\tau-\tau'). \end{aligned}\right\} \tag{16.9}$$

В соответствии с этими соотношениями выражение (16.7) для среднеквадратичной ошибки $\overline{e^2}$ можно преобразовать

к виду

$$\overline{e^2} = \int_0^\infty \int_0^\infty \{R_{ff}(\tau-\tau')[h(\tau)-h_1(\tau)][h(\tau')-h_1(\tau')] + R_{fn}(\tau-\tau')[h(\tau)-h_1(\tau)]h(\tau') + R_{nf}(\tau-\tau')h(\tau)[h(\tau')-h_1(\tau')] + R_{nn}(\tau-\tau')h(\tau)h(\tau')\}\,d\tau\,d\tau'. \tag{16.10}$$

Это соотношение служит для подсчета среднеквадратичной ошибки по функциям корреляции и реакции системы на единичный импульс.

Однако проектировщики систем регулирования предпочитают вести исследование, исходя непосредственно из передаточных функций $F(s)$ и $F_1(s)$. Чтобы следовать такому методу, мы используем изображения Фурье функций корреляции. Обозначим изображения Фурье функций корреляции через $\Phi_{ff}(\omega)$, $\Phi_{fn}(\omega)$, $\Phi_{nf}(\omega)$ и $\Phi_{nn}(\omega)$ соответственно, определив их равенствами

$$\left.\begin{aligned}
\Phi_{ff}(\omega) &= \frac{1}{\pi}\int_{-\infty}^{\infty} R_{ff}(\tau)\,e^{-i\omega\tau}\,d\tau,\\
\Phi_{fn}(\omega) &= \frac{1}{\pi}\int_{-\infty}^{\infty} R_{fn}(\tau)\,e^{-i\omega\tau}\,d\tau,\\
\Phi_{nf}(\omega) &= \frac{1}{\pi}\int_{-\infty}^{\infty} R_{nf}(\tau)\,e^{-i\omega\tau}\,d\tau,\\
\Phi_{nn}(\omega) &= \frac{1}{\pi}\int_{-\infty}^{\infty} R_{nn}(\tau)\,e^{-i\omega\tau}\,d\tau.
\end{aligned}\right\} \tag{16.11}$$

В силу симметричности функции $R_{ff}(t)$ по аргументу τ мы можем написать

$$\Phi_{ff}(\omega) = \frac{1}{\pi}\int_0^\infty R_{ff}(\tau)\left(e^{i\omega\tau}+e^{-i\omega\tau}\right)d\tau = \frac{2}{\pi}\int_0^\infty R_{ff}(\tau)\cos\omega\tau\,d\tau.$$

Сравнив это соотношение с соотношением (9.23), мы немедленно обнаружим, что функция $\Phi_{ff}(\omega)$ действительно

является спектральной плотностью сигнала $f(t)$. Подобным же образом функция $\Phi_{nn}(\omega)$ представляет собой спектральную плотность шума $n(t)$. Кроме того, так как функции взаимной корреляции связаны соотношением (16.9), легко убедиться, что подобная же связь существует и между их изображениями Фурье:

$$\left.\begin{aligned} \Phi_{fn}(-\omega) &= \Phi_{nf}(\omega), \\ \Phi_{nf}(-\omega) &= \Phi_{fn}(\omega). \end{aligned}\right\} \tag{16.12}$$

В силу теоремы об интеграле Фурье формулы обращения для уравнений (16.11) имеют вид

$$\left.\begin{aligned} R_{ff}(\tau-\tau') &= \frac{1}{2}\int_{-\infty}^{\infty} \Phi_{ff}(\omega)\, e^{i\omega(\tau-\tau')}\, d\omega, \\ R_{fn}(\tau-\tau') &= \frac{1}{2}\int_{-\infty}^{\infty} \Phi_{fn}(\omega)\, e^{i\omega(\tau-\tau')}\, d\omega, \\ R_{nf}(\tau-\tau') &= \frac{1}{2}\int_{-\infty}^{\infty} \Phi_{nf}(\omega)\, e^{i\omega(\tau-\tau')}\, d\omega, \\ R_{nn}(\tau-\tau') &= \frac{1}{2}\int_{-\infty}^{\infty} \Phi_{nn}(\omega)\, e^{i\omega(\tau-\tau')}\, d\omega. \end{aligned}\right\} \tag{16.13}$$

Подставляя (16.13) в (16.10), мы получим выражение среднеквадратичной ошибки через передаточные функции $F(s)$ и $F_1(s)$. Например, первое слагаемое правой части равенства (16.10) принимает вид

$$\begin{aligned} \frac{1}{2}\int_0^{\infty} d\tau \int_0^{\infty} d\tau' \int_{-\infty}^{\infty} d\omega\, \Phi_{ff}(\omega)\, e^{i\omega(\tau-\tau')} &\times \\ \times [h(\tau)-h_1(\tau)]\,[h(\tau')-h_1(\tau')] &= \\ = \frac{1}{2}\int_{-\infty}^{\infty} d\omega\, \Phi_{ff}(\omega) \int_0^{\infty} [h(\tau)-h_1(\tau)]\, e^{i\omega\tau}\, d\tau &\times \\ \times \int_0^{\infty} [h(\tau')-h_1(\tau')]\, e^{-i\omega\tau'}\, d\tau'&. \end{aligned}$$

Но функции $F(s)$ и $F_1(s)$ суть изображения по Лапласу функций $h(t)$ и $h_1(t)$, т. е.

$$\left.\begin{aligned} F(i\omega) &= \int_0^\infty h(t)\, e^{-i\omega t}\, dt, \\ F_1(i\omega) &= \int_0^\infty h_1(t)\, e^{-i\omega t}\, dt. \end{aligned}\right\} \tag{16.14}$$

Следовательно, это первое слагаемое можно переписать в виде

$$\frac{1}{2}\int_{-\infty}^{\infty} \Phi_{ff}(\omega)\,[F(i\omega) - F_1(i\omega)]\,[F(-i\omega) - F_1(-i\omega)]\, d\omega.$$

Подобным же образом можно преобразовать и другие слагаемые правой части равенства (16.10), и мы получим окончательно

$$\begin{aligned} \overline{e^2} = \frac{1}{2}\int_{-\infty}^{\infty} \{\Phi_{ff}(\omega)\,[F(i\omega) - F_1(i\omega)]\,[F(-i\omega) - F_1(-i\omega)] + \\ + \Phi_{fn}(\omega)\,[F(-i\omega) - F_1(-i\omega)]\,F(i\omega) + \\ + \Phi_{nf}(\omega)\,F(-i\omega)\,[F(i\omega) - F_1(i\omega)] + \\ + \Phi_{nn}(\omega)\,F(i\omega)\,F(-i\omega)\}\, d\omega. \end{aligned} \tag{16.15}$$

Подинтегральную функцию, заключенную в фигурные скобки, можно рассматривать как спектральную плотность ошибки $e(t)$. В силу соотношения (9.71), последнее слагаемое является спектральной плотностью шума, задержанного фильтром. Очевидно, что первое и четвертое слагаемые подинтегральной функции представляют собой вещественные величины, а второе и третье — комплексные. Однако на основании (16.12) эти комплексные величины — сопряженные, и поэтому их сумма является величиной вещественной.

16.2. Расчет оптимального фильтра по Филлипсу. Если статистические свойства сигнала и шума заданы через посредство различных функций корреляции, то изображе-

ния Фурье можно подсчитать с помощью соотношений (16.11). Тогда, коль скоро вид передаточной функции $F_1(s)$ установлен, для вычисления интеграла (16.15), определяющего среднеквадратичную ошибку, остается еще найти только передаточную функцию $F(s)$ фильтра. Поэтому оптимальным будет такой фильтр, передаточная функция которого будет минимизировать среднеквадратичную ошибку при заданных функциях Φ_{ff}, ... и $F_1(s)$.

Прямой метод решения задачи о расчете оптимального фильтра заключается в том, что мы задаем целесообразную форму функции $F(s)$, но пока оставляем неопределенными коэффициенты в выражении этой функции. Тогда после подстановки выбранного для $F(s)$ выражения в равенство (16.15) величину $\overline{e^2}$ можно выразить в виде функции от этих неопределенных постоянных.

В конечном счете постоянные определяются из условия, заключающегося в том, что среднеквадратичная ошибка должна достигать минимального значения на множестве этих постоянных. Таким образом, задача о расчете оптимального фильтра сводится к нахождению максимума или минимума некоторой известной функции. Именно этим путем шел Р. С. Филлипс[1]), разработавший такую теорию оптимального фильтра, выбрав передаточную функцию $F(s)$ в виде дробно-рациональной функции от s. Такое выражение для функции $F(s)$ является, конечно, естественным, вытекающим из нашего опыта в области линейных систем. Однако фактические вычисления по этому способу довольно трудоемки. По этой причине обычно предпочитают более изящную теорию Н. Винера и А. Н. Колмогорова, и мы не будем поэтому рассматривать далее теорию Филлипса.

16.3. Теория Винера—Колмогорова. Теория оптимального фильтра, развитая Н. Винером и А. Н. Колмогоровым, основана на решении вариационной задачи применительно к интегралу, находящемуся в правой части равенства (16.15). Пусть передаточная функция $F(s)$ уже характеризует оптимальный фильтр для данных функций Φ_{ff}, ... и $F_1(s)$. Тогда, образовав функцию $F(s)+\eta(s)$

---

[1]) Джеймс Х., Никольс Н., Филлипс Р., Теория следящих систем, М., 1951, гл. VII.

сравнения, близкую к $F(s)$, где $\eta(s)$ — произвольная вариация, и подставив эту функцию в равенство (16.15), найдем первую вариацию среднеквадратичной ошибки:

$$\begin{aligned}\delta\overline{e^2} = \frac{1}{2}\int_{-\infty}^{\infty} \eta(-i\omega)\{F(i\omega)[\Phi_{ff}(\omega) + \\ + \Phi_{fn}(\omega) + \Phi_{nf}(\omega) + \Phi_{nn}(\omega)] - \\ - F_1(i\omega)[\Phi_{ff}(\omega) + \Phi_{nf}(\omega)]\}\,d\omega + \\ + \frac{1}{2}\int_{-\infty}^{\infty} \eta(i\omega)\{F(-i\omega)\times \\ \times[\Phi_{ff}(\omega) + \Phi_{fn}(\omega) + \Phi_{nf}(\omega) + \Phi_{nn}(\omega)] - \\ - F_1(-i\omega)[\Phi_{ff}(\omega) + \Phi_{fn}(\omega)]\}\,d\omega. \qquad (16.16)\end{aligned}$$

Если функция $F(s)$ представляет собой передаточную функцию оптимального фильтра, то вариация $\delta\overline{e^2}$ должна обращаться в нуль при любой функции $\eta(s)$. Из этого условия мы можем вывести уравнение относительно функции $F(s)$. Однако, прежде чем получить это уравнение, нам придется совершить над уравнением (16.16) некоторые важные преобразования.

Прежде всего отметим, что спектральные плотности $\Phi_{ff}(\omega)$ и $\Phi_{nn}(\omega)$ суть симметрические функции ω. Принимая во внимание равенство (16.12), мы видим, что сумма функций Φ_{ff}, Φ_{fn}, Φ_{nf} и Φ_{nn} является четной функцией от ω. Поэтому мы имеем основание ожидать, что эту сумму Φ можно представить в виде произведения двух сопряженных множителей:

$$\Phi(\omega) = \Phi_{ff}(\omega) + \Phi_{fn}(\omega) + \Phi_{nf}(\omega) + \Phi_{nn}(\omega) = \Psi(i\omega)\,\Psi(-i\omega). \qquad (16.17)$$

По определению функции $\Psi(s)$ ее полюсы и нули лежат в левой полуплоскости плоскости комплексной переменной s, следовательно, полюсы и нули функции $\Psi(-s)$ лежат в правой полуплоскости этой плоскости. Но, в силу требований устойчивости, передаточная функция фильтра может иметь нули только в левой полуплоскости s. Таким образом, $F(s)$ и $\eta(s)$ суть функции, имеющие полюсы

только в левой полуплоскости s, а $F(-s)$ и $\eta(-s)$ — функции, имеющие полюсы только в правой полуплоскости.

Таким образом, класс функций, к которому должны принадлежать функции $F(s)$ и $\eta(s)$, ограничивается требованиями физического характера. Уяснив себе это обстоятельство, перепишем (16.16) в виде

$$\begin{aligned}\delta\overline{e^2} = \frac{1}{2}\int_{-\infty}^{\infty} \eta(-i\omega)\Psi(-i\omega)\times \\ \times\left\{F(i\omega)\Psi(i\omega) - \frac{F_1(i\omega)[\Phi_{ff}(\omega)+\Phi_{nf}(\omega)]}{\Psi(-i\omega)}\right\}d\omega + \\ + \frac{1}{2}\int_{-\infty}^{\infty} \eta(i\omega)\Psi(i\omega)\left\{F(-i\omega)\Psi(-i\omega) - \right. \\ \left. - \frac{F_1(-i\omega)[\Phi_{ff}(\omega)+\Phi_{fn}(\omega)]}{\Psi(i\omega)}\right\}d\omega. \quad (16.18)\end{aligned}$$

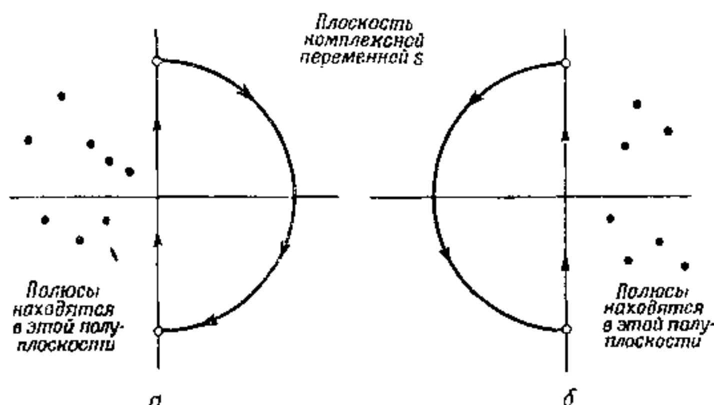
Однако не все из вторых слагаемых в фигурных скобках в правой части последнего равенства имеют важное значение. Пусть $H(s)$ и $K(s)$ суть функции с полюсами в левой полуплоскости комплексной переменной s, которые при больших s имеют порядок $1/s^n$, $n \geqslant 1$; тогда

$$\int_{-\infty}^{\infty} H(i\omega)K(i\omega)\,d\omega = i\oint H(s)K(s)\,ds,$$

где второй интеграл берется по замкнутому контуру, образованному мнимой осью и полукругом, охватывающим правую полуплоскость комплексной переменной s (фиг. 131, *а*). Но особые точки функций $H(s)$ и $K(s)$ лежат вне этого контура и, следовательно, величина этого интеграла равна нулю. При интегрировании произведения $H(-i\omega)\,K(-i\omega)$, особые точки которого лежат в правой полуплоскости s, контур интегрирования можно выбрать так, чтобы он охватывал левую полуплоскость s (фиг. 131, *б*). В этом случае опять-таки контур интегрирования не содержит особых точек и интеграл равен нулю. Но при интегрировании функций $H(i\omega)K(-i\omega)$ или

$H(-i\omega)K(i\omega)$ внутрь контура интегрирования уже попадут некоторые особые точки, и соответствующий интеграл будет отличен от нуля. Таким образом

$$\int_{\infty}^{\infty} H(i\omega)K(i\omega)\,d\omega = \int_{-\infty}^{\infty} H(-i\omega)K(-i\omega)\,d\omega = 0. \qquad (16.19)$$



Фиг. 131

Но

$$\left.\begin{aligned} \int_{-\infty}^{\infty} H(i\omega)K(-i\omega)\,d\omega \neq 0, \\ \int_{-\infty}^{\infty} H(-i\omega)K(i\omega)\,d\omega \neq 0. \end{aligned}\right\} \qquad (16.20)$$

Учитывая эти выводы, разобьем функцию

$$F_1(s)\,\Phi_{ff}(s/i)\,/\,\Psi(-s)$$

на две части. Одну из этих частей построим так, чтобы она имела особые точки в левой полуплоскости s, и обозначим ее символом $[\ ]_+$, а другую так, чтобы она имела особые точки в правой полуплоскости s, и обозначим ее

символом $[\;]_-$, т. е.

$$\frac{F_1(i\omega)[\Phi_{ff}(\omega)+\Phi_{nf}(\omega)]}{\Psi(-i\omega)}=$$
$$=\left[\frac{F_1(i\omega)\{\Phi_{ff}(\omega)+\Phi_{nf}(\omega)\}}{\Psi(-i\omega)}\right]_+ + \left[\frac{F_1(i\omega)\{\Phi_{ff}(\omega)+\Phi_{nf}(\omega)\}}{\Psi(-i\omega)}\right]_-. \tag{16.21}$$

Тогда, в силу равенств (16.19) и (16.20), равенство (16.18) можно переписать в виде

$$\delta\overline{e^2}=\frac{1}{2}\int_{-\infty}^{\infty}\eta(-i\omega)\,\Psi(-i\omega)\times$$
$$\times\left\{F(i\omega)\,\Psi(i\omega)-\left[\frac{F_1(i\omega)\{\Phi_{ff}(\omega)+\Phi_{nf}(\omega)\}}{\Psi(-i\omega)}\right]_+\right\}d\omega+$$
$$+\frac{1}{2}\int_{-\infty}^{\infty}\eta(i\omega)\,\Psi(i\omega)\times$$
$$\times\left\{F(-i\omega)\,\Psi(-i\omega)-\left[\frac{F_1(-i\omega)\{\Phi_{ff}(\omega)+\Phi_{fn}(\omega)\}}{\Psi(i\omega)}\right]_-\right\}d\omega. \tag{16.22}$$

Коль скоро функция $F(s)$ действительно является передаточной функцией оптимального фильтра, вариация $\delta\overline{e^2}$ должна уничтожаться при любых вариациях $\eta(s)$. Следовательно, величины, заключенные в фигурных скобках в правой части уравнения (16.22), должны уничтожаться.

Это условие определяет оптимальную передаточную функцию

$$F(s)=\frac{1}{\Psi(s)}\left[\frac{F_1(s)\{\Phi_{ff}(s/i)+\Phi_{nf}(s/i)\}}{\Psi(-s)}\right]_+. \tag{16.23}$$

Таково решение задачи об оптимальном фильтре, найденное Н. Винером и А. Н. Колмогоровым. Так как функция $\Psi(s)$ имеет нули только в левой полуплоскости s, то функция $F(s)$ будет иметь нули только в этой же левой полуплоскости s. Таким образом, $F(s)$ представляет собой устойчивую передаточную функцию. Операцию выделения из функции

$$F_1(s)\left\{\Phi_{ff}\left(\frac{s}{i}\right)+\Phi_{nf}\left(\frac{s}{i}\right)\right\}\Big/\Psi(-s)$$

части с полюсами в левой полуплоскости s можно выполнить также аналитически. В самом деле,

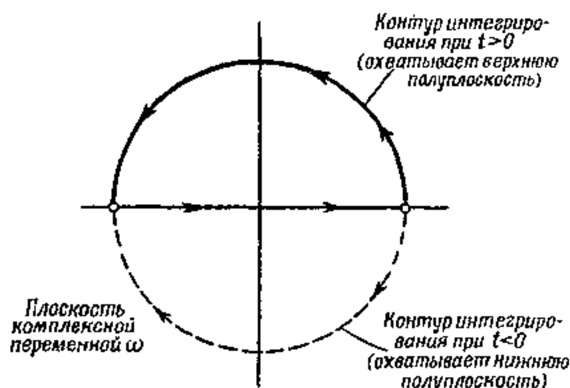
$$F(s) = \frac{1}{2\pi\Psi(s)} \int_{0}^{\infty} e^{-st} dt \int_{-\infty}^{\infty} \frac{F_1(i\omega)\{\Phi_{ff}(\omega) + \Phi_{nf}(\omega)\}}{\Psi(-i\omega)} e^{i\omega t} d\omega. \tag{16.24}$$

В справедливости этого соотношения легко убедиться путем интегрирования по определенному контуру на плоскости комплексной переменной ω. Выбор же из двух соотношений (16.23) и (16.24) того, которое используется для фактического подсчета оптимальной функции $F(s)$, зависит от конкретной задачи. Однако, как правило, гораздо легче исходить из соотношения (16.23). Во всех случаях свойства оптимального фильтра полностью определяются заданием передаточной функции $F_1(s)$ и спектральных плотностей Φ_{ff}, Φ_{fn}, Φ_{nf} и Φ_{nn} сигнала и шума. Когда шума нет, из (16.17) следует, что $\Phi_{nn} = \Phi_{fn} = \Phi_{nf} = 0$ и $\Phi_{ff}(\omega) = \Psi(i\omega)\Psi(-i\omega)$. Тогда, как и следовало ожидать, $F(s) = F_1(s)$, и ошибка $e(t)$ всегда равна нулю. При наличии шума функция $F(s)$ отлична от функции $F_1(s)$, и уничтожить среднеквадратичную ошибку нельзя даже с помощью наилучшего фильтра.

Операции выделения из некоторой функции части, имеющей полюсы в левой полуплоскости s, можно дать и другое истолкование. Как следует из преобразования (16.2), мы можем рассматривать $F(i\omega)$ как изображение Фурье реакции $h(t)$ на единичный импульс. Это соотношение доставляет также способ подсчета реакции $h(t)$ по частотной характеристике $F(i\omega)$. Но благодаря влиянию множителя $e^{i\omega t}$ в подинтегральной функции в правой части (16.2) контуры интегрирования на плоскости комплексной переменной ω имеют различную форму при положительных и отрицательных t, что изображено на фиг. 132. Если передаточная функция имеет полюсы только в левой полуплоскости плоскости комплексной переменной s, то функция $F(i\omega)$ имеет полюсы только в верхней полуплоскости комплексной переменной ω. Если $F(s)$ имеет полюсы только в правой полуплоскости s, то $F(i\omega)$ имеет полюсы только в нижней полуплоскости ω. Таким образом, при $t > 0$ контур интегрирования будет охватывать полюсы функ-

ции $F(i\omega)$ только в том случае, если передаточная функция $F(s)$ имеет полюсы в левой полуплоскости s; реакция $h(t)$ будет при этом отличной от нуля.

При $t < 0$ контур интегрирования не будет содержать ни одного полюса функции $F(i\omega)$ только в том случае, если все полюсы передаточной функции $F(s)$ лежат в левой полуплоскости s, и тогда $h(t) = 0$. С другой стороны,



Фиг. 132

если передаточная функция $F(s)$ имеет полюсы в правой полуплоскости s, что соответствует здесь случаю неустойчивой функции $F(s)$, то реакция $h(t)$ будет отличной от нуля даже при отрицательных t. Так как импульс подается в момент $t = 0$, то, говоря о реакции при отрицательных t, мы подразумеваем реакцию, наступающую до приложения импульса. Но никакая физическая система не может реагировать таким образом. Поэтому мы можем рассматривать операцию выделения из данной функции части, имеющей полюсы только в левой полуплоскости s, как операцию создания физически осуществимой передаточной функции, так что $h(t) = 0$ при $t < 0$. Боде и Шэннон [1]) дали объяснение решения уравнений (16.22) и (16.23), принадлежащего Винеру и Колмогорову, исходя из этого понятия физически осуществимой передаточной функции.

---

[1]) Bode H. W., Shannon C. E., Proc. IRE, 38, 417—425 (1950).

Остается выяснить еще одно обстоятельство. Одно из предположений, положенных в основу настоящей теории, предусматривает возможность представления суммы спектральных плотностей в виде произведения (16.17), что не всегда возможно даже для четной положительной функции $\Phi(\omega)$ от ω. Чтобы такое представление в виде произведения было возможно, функция $\Phi(\omega)$ должна удовлетворять условию Винера — Пэли [1]):

$$\int_{-\infty}^{\infty} \frac{|\ln \Phi(\omega)|}{1+\omega^2}\, d\omega < \infty. \tag{16.25}$$

В наших задачах функция $\Phi(\omega)$ или сводится к постоянной, в случае белого шума, или стремится к нулю при $\omega \to \infty$. Таким образом, критерий Винера — Пэли требует, чтобы стремление функции Φ к нулю при больших значениях ω протекало не слишком быстро. Этот критерий допускает функции типа ω^{-n}, однако при функциях, ведущих себя при больших ω подобно $e^{-|\omega|}$ или $e^{-\omega^2}$, интеграл расходится и соответствующую функцию $\Phi(\omega)$ уже нельзя представить в виде произведения двух комплексно-сопряженных функций. К счастью, спектры практически встречающихся сигналов и шума в общем случае являются дробно-рациональными функциями от ω^2. Поэтому представление спектральной плотности в виде произведения (16.17) обычно оказывается возможным.

16.4. Простые примеры. В качестве простого примера зададим следующее выражение спектральной плотности сигнала:

$$\Phi_{ff}(\omega) = \frac{1}{1+\omega^4}. \tag{16.26}$$

Что касается шума, то мы будем считать его белым. Тогда его спектральная плотность сводится к постоянной, т. е.

$$\Phi_{nn} = n^4. \tag{16.27}$$

---

[1]) Paley R. E. A. C., Wiener W., Fourier Transforms in the Complex Domain, Amer. Math. Soc. Colloquium Publication, vol. 19, p. 17, 1934.

Функции взаимной корреляции, следовательно, равны нулю, как и их изображения Фурье, т. е.

$$\Phi_{fn} = \Phi_{nf} = 0. \tag{16.28}$$

Пусть требуется построить оптимальный фильтр для дифференцирования сигнала, т. е. $F_1(s) = s$. Прежде всего заметим, что

$$\Phi(\omega) = \Phi_{ff}(\omega) + \Phi_{nn}(\omega) = \frac{(1+n^4) + n^4\omega^4}{1+\omega^4}.$$

Таким образом,

$$\Phi\left(\frac{s}{i}\right) = \Psi(s)\Psi(-s) = \frac{(1+n^4)+n^4 s^4}{1+s^4}.$$

Представление этой функции в виде произведения двух сопряженных функций мы можем получить, непосредственно используя ее вид. Принимая во внимание, что функция $\Psi(s)$ имеет нули и полюсы только в левой полуплоскости комплексной переменной s, мы можем немедленно записать $\Psi(s)$ в следующем виде:

$$\Psi(s) = \frac{n^2 s^2 + n\sqrt{2}\sqrt[4]{1+n^4}\,s + \sqrt{1+n^4}}{s^2+\sqrt{2}s+1}.$$

Тогда

$$F_1(s)\,\Phi_{ff}\left(\frac{s}{i}\right)\Big/\Psi(-s) =$$

$$= \frac{s}{(s^2+\sqrt{2}s+1)(n^2s^2 - n\sqrt{2}\sqrt[4]{1+n^4}\,s + \sqrt{1+n^4})} =$$

$$= \frac{As+B}{s^2+\sqrt{2}s+1} + \frac{Cs+D}{n^2s^2 - n\sqrt{2}\sqrt[4]{1+n^4}\,s+\sqrt{1+n^4}},$$

где A, B, C и D — постоянные. Очевидно, что часть, которая имеет полюсы только в левой полуплоскости s, определяется первым слагаемым, откуда следует

$$\left[\frac{F_1(s)\,\Phi_{ff}\left(\frac{s}{i}\right)}{\Psi(-s)}\right]_+ = \frac{As+B}{s^2+\sqrt{2}s+1}.$$

Подсчитав A и B, найдем окончательно

$$F(s) = \frac{1}{\Psi(s)}\left[\frac{F_1(s)\,\Phi_{ff}\left(\frac{s}{i}\right)}{\Psi(-s)}\right]_+ =$$
$$= \frac{1}{(n^2+\sqrt{1+n^4})\,(n+\sqrt[4]{1+n^4})}\times$$
$$\times\frac{(\sqrt[4]{1+n^4}-n)\,s-n\sqrt{2}}{n^2s^2+n\sqrt{2}\,\sqrt[4]{1+n^4}\,s+\sqrt{1+n^4}}\,. \qquad (16.29)$$

Таково выражение для передаточной функции оптимального фильтра. С уменьшением шума $n\to 0$ и функция $F(s)$ приближается к идеальному виду.

Рассмотрим другой интересный пример, соответствующий случаю очень высокого уровня шума при слабом полезном сигнале. Пусть шум будет опять белым, а его спектральная плотность Φ_{nn} равна

$$\Phi_{nn}(\omega) = 1. \qquad (16.30)$$

Здесь опять отсутствует взаимная корреляция между шумом и сигналом и равенства (16.28) остаются в силе. Пусть спектральная плотность сигнала задана выражением

$$\Phi_{ff}(\omega) = k\varphi(\omega), \qquad (16.31)$$

где k — малая величина, а $\varphi(\omega)$ — четная функция от ω. Обозначим через $K(s)$ часть функции $\varphi(s/i)$, которая имеет полюсы только в левой полуплоскости s, т. е.

$$K(s) = \left[\varphi\left(\frac{s}{i}\right)\right]_+ = \frac{1}{2\pi}\int_0^\infty e^{-st}\,dt\int_{-\infty}^\infty \varphi(\omega)\,e^{i\omega t}\,d\omega. \qquad (16.32)$$

Тогда, так как $\varphi(\omega)$ — четная функция от ω,

$$\varphi\left(\frac{s}{i}\right) = K(s)+K(-s) \qquad (16.33)$$

и

$$\Phi(\omega) = \Psi(i\omega)\,\Psi(-i\omega) = 1+k\varphi(\omega) \approx$$
$$\approx [1+kK(i\omega)]\,[1+kK(-i\omega)]. \qquad (16.34)$$

Поэтому, обозначив через $F_1(s)$ функцию, определяющую желаемое преобразование сигнала, передаточную функцию

оптимального фильтра найдем из соотношения

$$F(s) \approx \frac{k}{1+kK(s)}\left[\frac{F_1(s)\,\varphi\left(\frac{s}{i}\right)}{1+kK(-s)}\right]_+ . \qquad (16.35)$$

Мы получили здесь второе приближение при малых k. Первое же приближение еще проще:

$$F(s) \approx k\left[F_1(s)\varphi\left(\frac{s}{i}\right)\right]_+ . \qquad (16.36)$$

Конкретизируя рассматриваемый пример, зададим спектральную плотность сигнала в виде

$$\varphi(\omega) = \frac{1}{1+\omega^4}$$

и пусть $F_1(s) = s$; тогда при малых k

$$F(s) \approx -\frac{k}{2\sqrt{2}}\cdot\frac{1}{s^2+\sqrt{2}\,s+1} .$$

Этот результат служит подтверждением равенства (16.29): для того чтобы получить одно из этих соотношений из другого, нужно положить n достаточно большим и выбрать $k = 1/n^4$.

Таким образом, при сильном шумовом воздействии наблюдается весьма большое отклонение вида передаточной функции оптимального дифференцирующего фильтра от ее вида в идеальных условиях и эта оптимальная передаточная функция совершенно непохожа на функцию $F_1(s) = s$.

16.5. Приложения теории Винера—Колмогорова. Помимо простых примеров, рассмотренных в предыдущем пункте, теория Винера — Колмогорова имеет много очень важных приложений. В настоящем разделе мы остановимся на некоторых из них.

Упреждающие фильтры. На вход таких фильтров подается величина $x(t)$, представляющая собой сумму полезного сигнала $f(t)$ и шума $n(t)$. С выхода фильтра снимается величина $y(t)$, которая служит наилучшим при

ближением полезного сигнала не в момент t, а в момент $t+\alpha$, где α – положительное число. Таким образом

$$F_1(s)=e^{\alpha s}. \tag{16.37}$$

Зададим теперь сигнал в виде случайной функции переключения. Тогда, в силу соотношения (9.50), спектральная плотность этой функции выражается в виде

$$\Phi_{ff}=\frac{1}{1+\omega^2}. \tag{16.38}$$

Спектральная плотность шума в предположении, что шум белый, имеет вид

$$\Phi_{nn}=n^2. \tag{16.39}$$

Так как взаимная корреляция отсутствует, получаем

$$\Phi(\omega)=\Psi(i\omega)\Psi(-i\omega)=\frac{(1+n^2)+n^2\omega^2}{1+\omega^2}.$$

Поэтому

$$\Psi(i\omega)=\frac{\sqrt{1+n^2}+ni\omega}{1+i\omega},$$

откуда следует, что

$$\frac{F_1(i\omega)\Phi_{ff}(\omega)}{\Psi(-i\omega)}=\frac{e^{i\alpha\omega}}{(1+i\omega)(\sqrt{1+n^2}-ni\omega)}$$

В этом случае для подсчета передаточной функции $F(s)$ следует воспользоваться соотношением (16.24). Прежде всего, при $t>0$,

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{F_1(i\omega)\Phi_{ff}(\omega)}{\Psi(-i\omega)}e^{i\omega t}\,d\omega=$$

$$=\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{i\omega(t+\alpha)}\,d\omega}{(1+i\omega)(\sqrt{1+n^2}-ni\omega)}=\frac{e^{-(t+\alpha)}}{n+\sqrt{1+n^2}}.$$

Отсюда, учитывая (16.24), получаем

$$F(s) = \frac{1+s}{\sqrt{1+n^2}+ns}\int\limits_0^\infty e^{-st}\,\frac{e^{-(t+\alpha)}}{n+\sqrt{1+n^2}}\,dt =$$

$$= \frac{(1+s)\,e^{-\alpha}\int\limits_0^\infty e^{-(s+1)t}\,dt}{(n+\sqrt{1+n^2})(\sqrt{1+n^2}+ns)}.$$

Итак, в окончательном виде оптимальный упреждающий фильтр с временем α упреждения определяется передаточной функцией

$$F(s) = \frac{e^{-\alpha}}{(n+\sqrt{1+n^2})}\,\frac{1}{(\sqrt{1+n^2}+ns)}. \tag{16.40}$$

Фильтры запаздывания. Действие фильтра этого типа подобно действию упреждающего фильтра с той разницей, что в данном случае время α «упреждения» отрицательно. Но непосредственное приложение предыдущих преобразований к данному случаю не приводит к успеху.

Действительно, не существует линейной системы конечного порядка, которая бы в точности являлась оптимальным фильтром запаздывания. Более прямое решение можно построить, исходя из аппроксимации

$$F_1(s) = e^{\alpha s} \approx \left[\frac{1+\dfrac{\alpha s}{2\nu}}{1-\dfrac{\alpha s}{2\nu}}\right]^{\nu}, \qquad \alpha < 0 \tag{16.41}$$

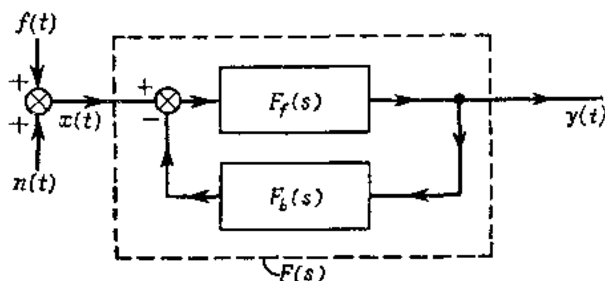
(ν — целое число). Таким способом мы придем к аппроксимации оптимального фильтра запаздывания.

Следящая система при наличии шума. Пусть $F_f(s)$ — передаточная функция прямой цепи, а $F_b(s)$ — передаточная функция цепи обратной связи следящей системы, схема которой изображена на фиг. 133. Пусть функция $F_1(s)$ определяет желаемое преобразование над сигналом. Задача заключается в нахождении оптимальной функции $F_b(s)$ при заданных функциях $F_f(s)$ и $F_1(s)$ и заданных свойствах сигнала и шума. Как вытекает из схемы системы,

показанной из фиг. 133, передаточная функция системы имеет вид

$$F(s) = \frac{F_f(s)}{1 + F_f(s) F_b(s)}$$

Но, согласно общей теории, оптимальная передаточная функция $F(s)$ определяется равенством (16.23) или (16.24). Зная $F(s)$, мы можем подсчитать оптимальную



Фиг. 133

передаточную функцию $F_b(s)$ цепи обратной связи с помощью выражения

$$F_b(s) = \frac{1}{F(s)} - \frac{1}{F_f(s)}. \qquad (16.42)$$

Следящие системы, содержащие источники шума. Выше мы исходили из предположения, что источник шума находится вне самой следящей системы и что сама следящая система не генерирует шумов. Однако очень часто шум генерируется внутри самой следящей системы. Например, система, схематически изображенная на фиг. 134, подвержена действию внешнего шума $n(t)$ и, кроме того, находится под влиянием внутреннего шума $m(t)$, исходящего от измерителя выхода или же представляющего собой возмущение на выходе. Пусть, как и выше, $F_f(s)$ — передаточная функция прямой цепи, $F_b(s)$ — передаточная функция цепи обратной связи и $F_1(s)$ — функция, определяющая желаемое преобразование над сигналом. Обозначим через $S(s)$, $N(s)$, $M(s)$, $Y(s)$ и $Z(s)$ изображения по Лапласу функций $f(t)$, $n(t)$, $m(t)$, $y(t)$ и $z(t)$ соответственно.

Тогда

$$F_f(s)\{S(s)+N(s)-F_b(s)[Y(s)+M(s)]\}=Y(s)$$

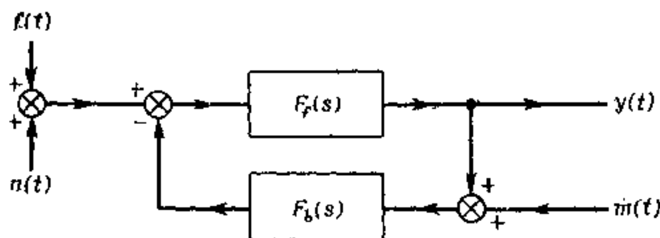
и

$$Z(s)=F_1(s)\,S(s).$$

Следовательно, изображение $E(s)$ по Лапласу ошибки $e(t)$ определяется выражением

$$E(s)=Y(s)-Z(s)=\frac{F_f(s)}{1+F_f(s)\,F_b(s)}[S(s)+N(s)]-$$

$$-\frac{F_f(s)\,F_b(s)}{1+F_f(s)\,F_b(s)}M(s)-F_1(s)\,S(s).$$



Фиг. 134

Обозначим

$$F(s)=\frac{F_f(s)\,F_b(s)}{1+F_f(s)\,F_b(s)}. \tag{16.43}$$

Тогда найдем, что

$$1-F(s)=\frac{1}{1+F_f(s)\,F_b(s)}$$

и

$$E(s)=[1-F(s)]\,[F_f(s)\,S(s)+F_f(s)\,N(s)-F_1(s)\,S(s)]-$$

$$-F(s)\,[M(s)+F_1(s)\,S(s)].$$

Это уравнение показывает, что задача о нахождении оптимальной передаточной функций $F(s)$ цепи обратной связи следящей системы равносильна задаче о построении фильтра, а изображение $S'(s)$ равносильного входного сигнала и изображение $N'(s)$ равносильного шума на входе опре-

деляются соотношениями

$$\left.\begin{aligned} S'(s) &= [F_f(s) - F_1(s)]\,S(s) + F_f(s)\,N(s), \\ N'(s) &= M(s) + F_1(s)\,S(s). \end{aligned}\right\} \qquad (16.44)$$

Таким образом, задача о разыскании оптимальной функции $F_b(s)$ свелась к задаче о разыскании оптимальной функции $F(s)$, причем равносильный сигнал и равносильный шум не зависят от искомой функции $F_b(s)$.

Допустим, что исходные сигнал и шумы независимы и существует только автокорреляция. Тогда статистические свойства определяются только спектральными плотностями Φ_{ff}, Φ_{nn} и Φ_{mm}. С помощью соотношений (16.44) подсчитаем спектральные плотности для задачи о нахождении равносильного фильтра

$$\begin{aligned} \Phi_{f'f'}\left(\frac{s}{i}\right) &= [F_f(s) - F_1(s)]\,[F_f(-s) - F_1(-s)]\,\Phi_{ff}\left(\frac{s}{i}\right) + \\ &\qquad + F_f(s)\,F_f(-s)\,\Phi_{nn}\left(\frac{s}{i}\right), \\ \Phi_{f'n'}\left(\frac{s}{i}\right) &= [F_f(-s) - F_1(-s)]\,F_1(s)\,\Phi_{ff}\left(\frac{s}{i}\right), \\ \Phi_{n'f'}\left(\frac{s}{i}\right) &= F_1(-s)\,[F_f(s) - F_1(s)]\,\Phi_{ff}\left(\frac{s}{i}\right), \qquad (16.45) \\ \Phi_{n'n'}\left(\frac{s}{i}\right) &= \Phi_{mm}\left(\frac{s}{i}\right) + F_1(s)\,F_1(-s)\,\Phi_{ff}\left(\frac{s}{i}\right). \end{aligned}$$

Таким образом, хотя в исходных сигнале и шумах взаимная корреляция отсутствует, в задаче о нахождении равносильного фильтра взаимная корреляция имеет место и определяется спектральными плотностями $\Phi_{f'n'}$ и $\Phi_{n'f'}$. Следовательно, функцией, которую мы должны представить в виде произведения комплексно-сопряженных функций, является, в силу (16.45), функция

$$\begin{aligned} \Phi(\omega) &= \Psi(i\omega)\,\Psi(-i\omega) = \\ &= F_f(i\omega)\,F_f(-i\omega)\,[\Phi_{ff}(\omega) + \Phi_{nn}(\omega)] + \Phi_{mm}(\omega). \qquad (16.46) \end{aligned}$$

Поэтому на основании равенства (16.23) оптимальная

функция $F(s)$ выражается соотношением

$$F(s) = \frac{1}{\Psi^{+}(s)} \times$$

$$\times \left[\frac{F_f(s) F_f(-s) \left\{ \Phi_{ff}\left(\frac{s}{i}\right) + \Phi_{nn}\left(\frac{s}{i}\right) \right\} - F_i(s) F_i(-s) \Phi_{ff}\left(\frac{s}{i}\right)}{\Psi(-s)} \right]_+ . \tag{16.47}$$

Как только функция $F(s)$ найдена, с помощью соотношения (16.43) определяется и оптимальная передаточная функция $F_b(s)$ цепи обратной связи:

$$F_b(s) = \frac{1/F_f(s)}{[1/F(s)] - 1} . \tag{16.48}$$

Принудительное ограничение мощности. Рассмотрим следящую систему, схематически изображенную на фиг. 135 и предназначенную для возможно более точного воспроизведения выходной координатой $y(t)$ входной координаты $f(t) = x(t)$. Пусть $F_a(s)$ — передаточная функция усилителя, а $F_m(s)$ — передаточная функция исполнительного привода. Если технические условия на разработку системы требуют обеспечения возможно меньшей величины среднего квадрата ошибки $e(t) = y(t) - f(t)$ и разрешают изменять с этой целью передаточную функцию $F_a(s)$, то во время действия системы мощность управляющего сигнала, подаваемого на вход исполнительного привода, может достичь очень большой величины. Во избежание этого явления мы должны потребовать, чтобы средняя мощность на входе двигателя равнялась заранее заданной величине. В данном случае среднеквадратичная ошибка определяется выражением

$$\overline{e^2} = \frac{1}{2} \int_{-\infty}^{\infty} \left[\frac{F_a(i\omega) F_m(i\omega)}{1 + F_a(i\omega) F_m(i\omega)} - 1 \right] \times$$

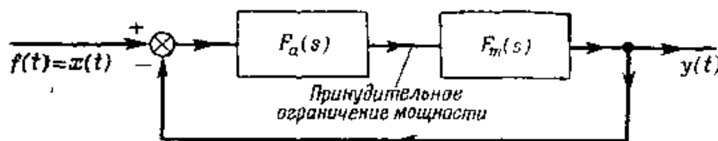
$$\times \left[\frac{F_a(-i\omega) F_m(-i\omega)}{1 + F_a(-i\omega) F_m(-i\omega)} - 1 \right] \Phi_{ff}(\omega)\, d\omega. \tag{16.49}$$

Средняя мощность входного сигнала исполнительного двигателя характеризуется среднеквадратичным значением управляющего сигнала этого двигателя; эту среднюю

мощность надо поддерживать равной предписанной величине σ^2. Таким образом,

$$\sigma^2 = \frac{1}{2} \int_{-\infty}^{\infty} \left| \frac{F_a(i\omega)}{1 + F_a(i\omega) F_m(i\omega)} \right|^2 \Phi_{ff}(\omega)\, d\omega. \qquad (16.50)$$

С помощью метода неопределенных множителей Лагранжа задачу о разыскании минимума $\overline{e^2}$ при условии, выражаемом соотношением (16.50), можно свести к задаче



Фиг. 135

о разыскании минимума функции $\overline{e^2} + \lambda\sigma^2$, где λ — постоянная усилителя. Тогда задача приводится к нахождению минимума интеграла

$$\overline{e^2} + \lambda\sigma^2 = \frac{1}{2} \int_{-\infty}^{\infty} \Big\{ [F(i\omega) - 1][F(-i\omega) - 1]\, \Phi_{ff}(\omega) + \\ + F(i\omega) F(-i\omega) \frac{\lambda \Phi_{ff}(\omega)}{F_m(i\omega) F_m(-i\omega)} \Big\}\, d\omega, \qquad (16.51)$$

где

$$F(s) = \frac{F_a(s) F_m(s)}{1 + F_a(s) F_m(s)}. \qquad (16.52)$$

Сравнивая (16.51) с интегралом, находящимся в правой части соотношения (16.15), мы убедимся, что данная задача равносильна задаче о фильтре при $F_1(s) = 1$, $\Phi_{fn} = \Phi_{nf} = 0$, и при спектральной плотности равносильного шума

$$\Phi_{nn}(\omega) = \frac{\lambda \Phi_{ff}(\omega)}{F_m(i\omega) F_m(-i\omega)}.$$

Представление спектральной плотности $\Phi(\omega)$ в форме произведения двух комплексно-сопряженных функций имеет здесь вид

$$\Phi(\omega) = \Psi(i\omega)\,\Psi(-i\omega) = \left[1 + \frac{\lambda}{F_m(i\omega)\,F_m(-i\omega)}\right]\Phi_{ff}(\omega). \qquad (16.53)$$

С помощью равенства (16.23) мы придем к следующему выражению оптимальной функции $F(s)$:

$$F(s) = \frac{1}{\Psi(s)}\left[\frac{\Phi_{ff}\left(\frac{s}{i}\right)}{\Psi(-s)}\right]_{+}. \qquad (16.54)$$

Соотношения (16.52) и (16.54) определяют передаточную функцию $F_a(s)$ оптимального усилителя с точностью до постоянной λ. Эта постоянная определяется с помощью величины σ^2, ограничивающей его выходную мощность, в соответствии с равенством (16.50).

Мы изложили несколько задач, на которых, возможно, удалось продемонстрировать широту проблем, решаемых применением теории Винера — Колмогорова. Целесообразно отметить, что с помощью этой теории можно успешно решить задачу о разработке следящих систем при случайных входах, поставленную в п.9.12. Такая задача в области следящих систем служит еще одним примером синтеза по заданным техническим условиям, что в самом общем виде рассматривалось в гл.XIV. Коль скоро имеется возможность установить эти технические условия, то тем самым полностью определяются и характеристики оптимальной системы. Кроме того, построенная таким образом система регулирования оказывается линейной системой с постоянными коэффициентами, что объясняется специфической особенностью выбранных технических условий и свойствами объекта регулирования. Таким образом, это понятие более специфической задачи синтеза следящих систем является шагом вперед по сравнению с основными принципами синтеза систем с обратной связью, рассмотренными в предыдущих главах. Можно, повидимому, считать, что на это конкретное приложение теории фильтров впервые указали Боксенбом и Новик [1]).

[1]) Boksenbom A. S., Novik D., NACA TN 2939 (1953).

16.6. Оптимальный фильтр, отделяющий полезный сигнал от шума. Во многих системах регулирования возникает задача о выделении полезного сигнала $f(t)$ в условиях сильной помехи со стороны шумов. В таких задачах общая форма сигнала обычно известна; установить же требуется момент t_0 времени, в который сигнал достигает ожидаемой величины $f(t_0)$. Например, в случае радиолокационной системы нам известно, что сигнал представляет собой импульс определенной формы. Задача состоит в определении момента времени, в который этот сигнал достигает наибольшей интенсивности. Если это наибольшее значение сигнала имеет место в момент времени t_0, то фильтр даст наименьшее искажение именно в момент t_0. Поэтому оптимальный фильтр нужно спроектировать таким образом, чтобы величина $f_o(t)$ сигнала, прошедшего через фильтр, равнялась в момент t_0 величине $f(t_0)$. Следовательно, связь, налагаемая на фильтр, определяется соотношением

$$f_o(t_0) = f(t_0) = \text{const}. \tag{16.55}$$

Пусть $n(t)$ — шум на входе системы, а $n_o(t)$ — соответствующий ему выход. При введении фильтра мы ставим своей целью в возможно большей степени уменьшить шум и, таким образом, можем написать

$$\overline{n_o^2(t)} = \min. \tag{16.56}$$

Теперь задача заключается в определении передаточной функции $F(s)$ фильтра или равносильного устройства и реакции $h(t)$ фильтра на единичный импульс при заданных $f(t)$, t_0 и характеристиках R_{nn} или Φ_{nn} шума при условии

$$\overline{n_o^2(t)} - 2\lambda f_o(t_0) = \min, \tag{16.57}$$

где λ — множитель Лагранжа. В общей форме эту задачу решили Задэ и Рагаццини[1]; перейдем к изложению их исследования.

Допустим, что функцию $\Phi_{nn}(\omega)$ можно представить в виде произведения

$$\Phi_{nn}(\omega) = \Psi(i\omega)\Psi(-i\omega), \tag{16.58}$$

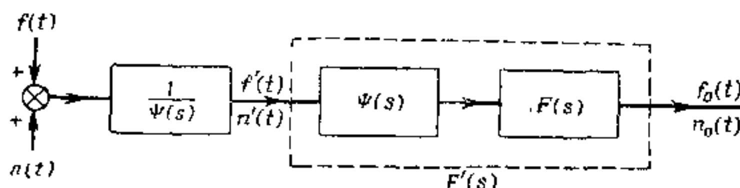
---

[1]) Zadeh L. A., Ragazzini I. R. Proc. I. R. E., 40, 1223—1231 (1952).

где функция $\Psi(s)$ имеет нули и полюсы только в левой полуплоскости комплексной переменной s. В связи с этим удобнее рассматривать усилитель в виде последовательного соединения двух усилителей, причем первый из них обладает передаточной функцией $1/\Psi(s)$, а второй – передаточной функцией $\Psi(s)$, и оба этих усилителя последовательно соединены с фильтром, как показано на фиг. 136. Эта система равносильна исходной системе. Последние две передаточные функции объединим в одну передаточную функцию $F'(s)$:

$$F'(s) = \Psi(s) F(s), \tag{16.59}$$

которой отвечает реакция $h'(t)$ на единичный импульс. Входом блока, описываемого передаточной функцией $F'(s)$,



Фиг. 136

служат полезный сигнал $f'(t)$ и шум $n'(t)$. Обозначив через $S(i\omega)$ изображение Фурье сигнала $f(t)$, т. е.

$$S(i\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t}\, dt, \tag{16.60}$$

найдем, что

$$f'(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{S(i\omega)}{\Psi(i\omega)} e^{i\omega t}\, d\omega. \tag{16.61}$$

В этих условиях спектральная плотность функции $n'(t)$, в силу (9.71), определяется соотношением

$$\frac{\Phi_{nn}(\omega)}{\Psi(i\omega)\Psi(-i\omega)} = 1.$$

Поэтому шум является теперь белым шумом с автокорреляционной функцией

$$R_{n'n'}(\tau) = \delta(\tau). \tag{16.62}$$

Соответственно выходы системы суть

$$f_o(t)=\int_0^\infty h'(\tau)f'(t-\tau)\,d\tau \tag{16.63}$$

и

$$n_o(t)=\int_0^\infty h'(\tau)\,n'(t-\tau)\,d\tau. \tag{16.64}$$

Таким образом, среднеквадратичное значение шума на выходе равно среднему по множеству функции $n_o^2(t)$ или

$$\overline{n_o^2(t)}=\int_0^\infty d\tau\int_0^\infty d\tau' h'(\tau)\,h'(\tau')\,\overline{n'(t-\tau)\,n'(t-\tau')}=$$

$$=\int_0^\infty\int_0^\infty h'(\tau)\,h'(\tau')\,R_{n'n'}(\tau-\tau')\,d\tau\,d\tau'.$$

Но функция корреляции шума $n'(t)$ определена соотношением (16.62); следовательно

$$\overline{n_o^2}=\int_0^\infty [h'(t)]^2\,dt. \tag{16.65}$$

Поэтому с учетом равенств (16.63) и (16.65) условие (16.57) можно преобразовать к виду

$$\int_0^\infty [h'(t)]^2\,dt-2\lambda\int_0^\infty h'(t)f'(t_0-t)\,dt=\min.$$

Это условие можно переписать в виде

$$\int_0^\infty [h'(t)-\lambda f'(t_0-t)]^2\,dt-\lambda^2\int_0^\infty [f'(t_0-t)]^2\,dt=\min. \tag{16.66}$$

Но, коль скоро функции $f(t)$ и $\Phi_{nn}(t)$ заданы по условиям задачи, функция $f'(t)$ становится определенной; ее можно вычислить с помощью равенства (16.61). Поэтому

второй интеграл в левой части (16.66) сводится к постоянной. Первый же интеграл положителен или равен нулю. Чтобы сумма в левой части равенства (16.66) достигла минимального значения, первый интеграл этой суммы должен обратиться в нуль. Но этот интеграл обращается в нуль только в том случае, если исчезает величина, заключенная в квадратных скобках в выражении подинтегральной функции, т. е. если

$$h'(t) = \lambda f'(t_0 - t), \qquad t > 0. \tag{16.67}$$

Естественно, что в физически осуществимых системах $h'(t) \equiv 0$ при $t < 0$. Другими словами, *оптимальная реакция* F' *на импульс совпадает для положительных* t *с изображением* $f'(t)$ *относительно* $t_0/2$. Этот результат для белого шума был впервые получен Д. О. Нортом.

Основываясь на результате, выражаемом равенством (16.67), можно сразу же выписать передаточную функцию $F(s)$ оптимального фильтра исходной задачи, учитывая соотношения (16.59) и (16.61):

$$F(s) = \frac{\lambda}{2\pi\Psi(s)} \int_0^\infty dt\, e^{-st} \int_0^\infty \frac{S(-i\omega)}{\Psi(-i\omega)} e^{i\omega(t-t_0)}\, d\omega. \tag{16.68}$$

Постоянная λ, входящая в последнее равенство, в конечном счете определяется заданием постоянной $f(t_0)$ (16.55). Полученный фильтр весьма похож на оптимальный фильтр Винера, определенный выражением (16.24). В самом деле, мы можем написать

$$F(s) = \frac{\lambda}{\Psi(s)} \left[\frac{e^{-st_0} S(-s)}{\Psi(-s)} \right]_+, \tag{16.69}$$

где символ $[\ ]_+$ попрежнему означает операцию взятия той части функции, заключенной в квадратных скобках, которая имеет полюсы только в левой полуплоскости S, т. е. операцию превращения передаточной функции в функцию, физически осуществимую. Выражения (16.68) и (16.69), полученные Задэ и Рагаццини, определяют оптимальный фильтр для выделения полезного сигнала. Благодаря применению искусственного приема, с помощью которого задача сводится к равносильной, согласно схеме

на фиг. 136, вывод этих выражений для случая белого шума значительно упрощается. Такой прием, вообще говоря, очень полезен в том отношении, что с его помощью удается упрощать исследование сложных задач оптимального регулирования.

16.7. Другие оптимальные фильтры. Одно из основных предположений, на котором построена теория фильтров, разобранная в предыдущих разделах, заключается в стационарном характере рассматриваемых случайных функций: как полезного сигнала, так и шума. Эти функции не могут быть подлинно стационарными в течение очень длительного времени; их стационарный характер будет нарушаться за счет естественного изменения состояния системы или же при преднамеренных изменениях ее рабочего режима. Вернее предполагать, что случайные входы сохраняют стационарный характер только в ограниченном интервале T времени. В интервалах же времени, более длинных по сравнению с T, случайные функции не остаются стационарными. Поэтому если фильтр рассчитан, исходя из допущения о стационарном характере случайных функций, и если характеристическое время фильтра больше T, то действительность не будет отвечать теоретическим посылкам и качество действия фильтра будет более низким. В таких случаях лучше отступить от теоретического «оптимального» фильтра и применять фильтр с меньшим характеристическим временем.

Более удачное решение задачи основано на построении теории оптимального фильтра с учетом времени T. Для этого мы потребуем, чтобы реакция $h(t)$ фильтра на единичный импульс уничтожалась вне интервала $0 \leqslant t \leqslant T$. В этом случае выход $y(t)$ выражается через вход $x(t)$ посредством соотношения

$$y(t) = \int_0^T h(\tau)\, x(t-\tau)\, d\tau. \tag{16.70}$$

Как показывает это соотношение, выход зависит от входа только на протяжении конечного промежутка времени T, отсчитываемого в обратном направлении от данного момента. Поэтому такие фильтры можно назвать *фильтрами с конеч-*

ной памятью. Фильтры же, рассмотренные в предыдущих разделах, для которых $T \to \infty$, являются, таким образом, *фильтрами с бесконечной памятью.*

Фильтры с конечной памятью рассматривали Задэ и Рагаццини. Им же[1]) принадлежит решение задачи об оптимальном фильтре для задачи выделения полезного сигнала, подобной задаче, рассмотренной в предыдущем разделе. Они решили[2]) также задачу о синтезе оптимального фильтра в более сложном случае, в котором входной сигнал состоит из двух частей: стационарной случайной функции и неслучайной функции, представимой с помощью многочлена n-й степени от t. Эта задача о синтезе фильтра конечной памяти решается на основании следующих критериев качества: мы требуем, во-первых, уничтожения средней ошибки, а во-вторых, —уничтожения среднеквадратичной ошибки. Здесь уничтожение средней ошибки уже не обеспечивается автоматически из-за влияния неслучайной части сигнала. Однако данные ими решения этих задач в общем случае трудно физически осуществить с помощью простых цепей типа RC. В самом деле, даже для более простой задачи фильтров с бесконечной памятью иногда бывает затруднительно осуществить решение, определяемое соотношениями (16.68) и (16.69).

Фильтры, применяемые на практике, могут быть лишь аппроксимацией фильтров, обеспечивающих теоретический оптимум. Следовательно, ценность теоретического решения состоит главным образом в выяснении направления, в котором следует вести проектирование, и в определении свойств идеального образца.

16.8. Общая задача фильтрации. Разумеется, при отказе от неподходящих в общем случае цепей типа RC и при использовании в качестве фильтров моделирующих устройств и даже цифровых машин мы все же сумеем применить на практике сложный процесс синтеза теоретических оптимальных фильтров. В этом случае удастся практически достичь теоретического оптимального результата.

---

[1]) См. примечание на стр. 389.

[2]) Zadeh L. A., Ragazzini J. R., Journ. Appl. Phys., 21, 645—654 (1950).

Однако с привлечением электромеханического вычислительного устройства как составной части системы фильтра в большой степени возрастает сложность всей совокупной системы, что может оправдываться только в самых крайних случаях. Если система получается очень усложненной и дорогостоящей, то уместен вопрос, действительно ли мы добились самого лучшего качества. Пути получения оптимального качества, выявленные с помощью теоретических рассуждений, проведенных в предыдущих разделах, приводят к оптимальной системе фильтра лишь в пределах ограничений, связанных с допущениями, на которых основана эта теория. Например, из этой теории следует, что два случайных сигнала, обладающих одинаковой функцией корреляции или одинаковой спектральной плотностью, требуют одного и того же оптимального фильтра, что в некотором смысле связано с известной общностью технических условий.

Конечно, если бы мы располагали бо́льшими статистическими данными о сигнале, а не только знали его спектральную плотность, мы были бы в состоянии различать эти два сигнала и смогли бы улучшить конструкцию оптимального фильтра путем использования таких дополнительных данных. Тогда мы смогли бы достичь качества, даже лучшего по сравнению с тем, которое возможно при так называемом «оптимальном» фильтре. Очевидно, что такой обобщенный подход к задаче о синтезе фильтра непременно потребует и более разработанной теории вероятностей, чем та, которой мы пользовались. Здесь может иметь важные приложения и недавно сложившаяся наука—теория информации. В этой области начало было заложено таким «вероятностным» направлением в задаче об отделении полезного сигнала от шума[1]). Но многое остается еще сделать.

Теория Винера—Колмогорова об оптимальном фильтре основана на критерии среднеквадратичной ошибки. Применяя этот критерий, мы делаем существенный упор на уменьшение до минимума больших ошибок, не уделяя значительного внимания малым ошибкам. Однако во многих случаях мы больше заинтересованы в возможно более зна-

---

[1]) См., например, Woodward P. M., Davies I. L., Phil. Mag., **41**, 1001—1017 (1950); Proc. IRE, **39**, 1521—1524 (1951); Journ. Inst. Elec. Engs. London, **99** (3), 37—51 (1952); Slattery T. G., Proc. IRE, **40**, 1232—1236 (1952).

чительном уменьшении часто повторяющихся ошибок и нас в меньшей степени обеспокоят большие, но не частые ошибки.

Возможно также, что функция распределения вероятностей будет иметь резко несимметричный характер и среднее значение будет в значительной степени смещено по отношению к максимуму этой функции.

В таких случаях критерий среднеквадратичной ошибки совершенно непригоден. В качестве простого примера обратимся к задаче Боде и Шэннона[1]) о предвидении того, будет ли на следующий день хорошая погода. Поскольку ясные дни составляют большинство и не существует дней, которые бы знаменовались отрицательными осадками, способными компенсировать дни с осадками, функция распределения вероятностей имеет весьма несимметричный характер. На основании этой функции среднее значение, определяемое исходя из минимума среднеквадратичной ошибки, может соответствовать дню с моросящим дождем. Однако для лица, организующего пикник, такое предсказание не имеет никакой ценности. Оно заинтересовано в том, чтобы знать вероятность наступления дня с настоящей ясной погодой, ибо даже малое количество осадков испортит весь пикник.

Теория оптимального фильтра, рассмотренная в этой главе, также исходит из допущения о линейности этого фильтра в том смысле, что дифференциальное уравнение, связывающее вход фильтра с его выходом, представляет собой линейное дифференциальное уравнение с постоянными коэффициентами. Но это ограничение имеет явно произвольный характер, и оно наложено только с целью упрощения теории и в расчете на то, что такие фильтры легко синтезировать с помощью контуров типа RC. Для систем регулирования с переменными параметрами, как, например, для системы регулирования в задаче о наведении ракеты на цель (гл. XIII), такие фильтры, конечно, непригодны. В случае систем с переменными параметрами надлежащий фильтр также должен обладать переменными параметрами. Если фильтр остается все же линейным, так что к нему применим принцип наложения, то соотношение между входом и выхо-

[1]) Bode N. W. Shannon C. E., работа, цит. на стр. 376.

дом будет попрежнему определяться реакцией на единичный импульс. Но в этом случае эта реакция уже будет функцией h двух переменных t и t^*. Переменная t определяет момент времени съема реакции, а t^*—момент времени приложения импульса.

Тогда подсчет выходной величины $y(t)$ по входной величине $x(t)$ производится с помощью соотношения

$$y(t) = \int_0^\infty h(t, t - \tau)\, x(t - \tau)\, d\tau. \tag{16.71}$$

Следовательно, задача о построении оптимального фильтра заключается прежде всего в определении функции $h(t, \tau)$ и затем в нахождении способа физического осуществления этой оптимальной реакции на единичный импульс.

Глава XVII

УЛЬТРАУСТОЙЧИВОСТЬ И МУЛЬТИУСТОЙЧИВОСТЬ

В предыдущих главах мы рассмотрели методы проектирования весьма сложных систем автоматического регулирования, позволяющие удовлетворить почти любым наперед заданным техническим условиям. Конечно, чем сложнее система, тем больше вероятность неправильного срабатывания системы из-за ошибок в сборке ее или из-за выхода из строя отдельных элементов системы. Поэтому в сложных системах автоматического регулирования исключительное значение преобретает задача обеспечения надежности конструкции в эксплоатационных условиях. В настоящей и в последующей главах мы рассмотрим эту задачу с двух различных точек зрения.

Настоящая глава посвящена вопросу о возможности придания системе определенной степени гибкости и способности к приспособлению, так чтобы система регулирования сама, без вмешательства человека, автоматически выправляла случайные и непредвиденные ошибки, допущенные при ее синтезе. Таким образом, подобная система автоматического регулирования способна «сознавать», как ей следует себя вести, и обладает почти *гомеостатическим механизмом* живых существ, благодаря которому они выживают при различных меняющихся условиях внешней среды[1]). Естественно, что это понятие самонастройки в сложной системе заимствовано из изучения поведения живых существ, так как именно у живых существ характеристики этой самонастройки выступают в наиболее отчетливом виде. Наши рассуждения в этой главе основаны на замечательной книге У. Р. Эшби[2]) о происхождении единственной в своем роде

---

[1]) Конечно, слово «сознавать» в данном случае следует взять в кавычки.—*Прим. ред.*

[2]) Ashby W. R., Design for a Brain, New York, 1952.

способности нервной системы вызывать поведение живого существа, отличающееся характером приспособления. Могут существовать различные мнения о том, как в действительности построен мозг того или иного живого существа. Но наша цель состоит просто в том, чтобы показать, что с помощью механических средств можно заставить систему приспосабливаться. При этом мы не рассматриваем здесь вопроса о том, является ли предлагаемое механическое решение задачи единственным или нет.

17.1. Ультраустойчивые системы. Для простоты рассмотрим некоторую *автономную* систему, определяющуюся двумя переменными y_1 и y_2, как это уже встречалось в п.10.5. Фазовой плоскостью служит, следовательно, плоскость переменных y_1, y_2. Обозначая через t время, напишем систему двух дифференциальных уравнений, определяющих поведение системы, в виде

$$\left.\begin{aligned} \frac{dy_1}{dt} &= f_1(y_1, y_2; \zeta), \\ \frac{dy_2}{dt} &= f_2(y_1, y_2; \zeta), \end{aligned}\right\} \tag{17.1}$$

где под знаками функций f_1 и f_2 включен дополнительный параметр ζ. Это означает, что функциональное соотношение, связывающее dy_1/dt и dy_2/dt, с одной стороны, и y_1 и y_2 — с другой, определено только в том случае, когда задано значение параметра ζ. В частности, пусть ζ пробегает дискретную последовательность значений. При этом поведение системы определяется геометрическими местами точек (y_1, y_2) (фазовыми траекториями), причем точка (y_1, y_2) движется по соответствующей фазовой траектории в направлении, отвечающем возрастанию времени; каждое такое геометрическое место начинается в различных начальных точках на фазовой плоскости.

В этих условиях очевидно, что число различных форм поведения системы равно числу различных значений этого параметра ζ, причем, по условию, ζ принимает значения, принадлежащие дискретному множеству. Пусть, например, от параметра ζ зависит линейная система, рассмотренная в п.10.5 (причем $y_1 = y$ и $y_2 = \dot{y}$), а ζ может принимать

пять различных значений: отрицательное, меньшее -1; отрицательное, заключенное между 0 и -1; равное нулю; положительное, заключенное между нулем и единицей, и, наконец, положительное, большее 1. Этим пяти значениям ζ будут отвечать пять видов процессов в системе, графически изображенных на фиг. 82—86. Другой пример доставляет система уравнений

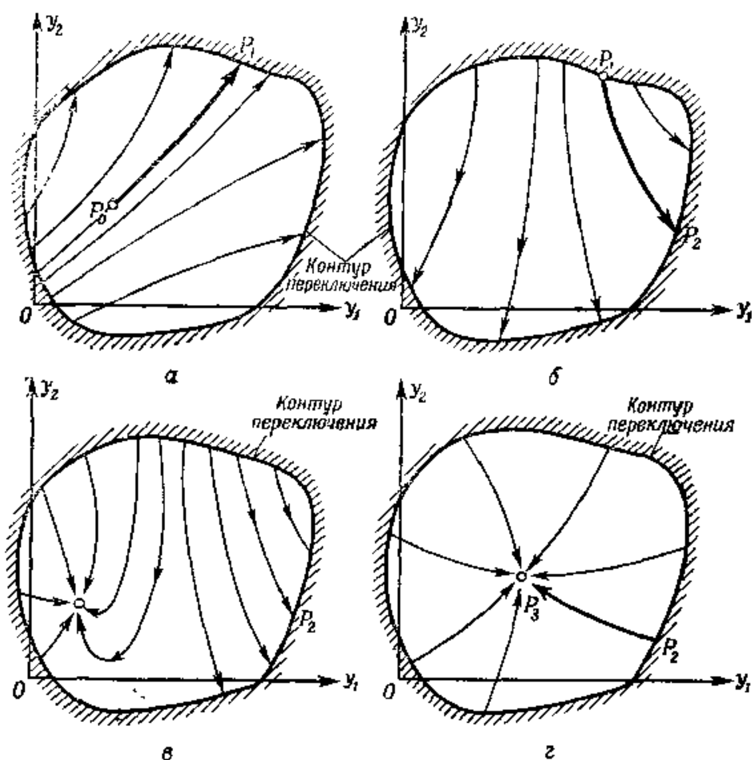
$$\left.\begin{aligned} \frac{dy_1}{dt} &= a_{11}(\zeta)\, y_1 + a_{12}(\zeta)\, y_2, \\ \frac{dy_2}{dt} &= a_{21}(\zeta)\, y_1 + a_{22}(\zeta)\, y_2, \end{aligned}\right\} \tag{17.2}$$

где коэффициенты a_{11}, a_{12}, a_{21} и a_{22} суть монотонные функции параметра ζ. В этих условиях существует столько же различных сочетаний этих коэффициентов, сколько имеется различных значений параметра ζ. Каждое такое сочетание коэффициентов определяет течение процессов в системе.

В некоторых случаях такие системы будут устойчивы в том смысле, что все фазовые траектории будут сходиться к некоторой точке на фазовой плоскости — точке устойчивого равновесия. В других случаях эти системы будут неустойчивыми в том смысле, что фазовые траектории будут удаляться от точек равновесия.

Конечно, чтобы процессы в системе протекали удовлетворительно, система должна быть устойчивой. Мы создадим желаемую способность системы к приспособлению, если сумеем придать ей свойство автоматически отбрасывать неустойчивое течение процессов и сохранять состояния, в которых эти процессы протекают устойчиво. Поступим следующим образом: окружим желаемую точку равновесия системы замкнутым контуром и введем в систему переключающее устройство, которое будет срабатывать так, что значение параметра ζ скачкообразно изменяется каждый раз, когда фазовая траектория пересекает границу этого контура. Исследуем действие такой схемы. Обращаясь к исследованию процесса, начинающегося в точке P_0 на фиг. 137, *а*, мы видим, что этот процесс будет неустойчивым и соответствующая фазовая траектория пересечет контур переключения в некоторой точке, которую обозначим через P_1

В момент пересечения переключатель сработает таким образом, что значение параметра ζ изменится на конечную величину, в результате чего характер процесса в системе станет иным и будет определяться фазовой траекторией, показанной на фиг. 137, *б*. Новый процесс оказывается здесь



Фиг. 137

также неустойчивым, и поэтому состояние системы сместится из точки P_1 в точку P_2, в которой фазовая траектория вновь пересечет контур переключения. Опять сработает переключающее устройство, придавая параметру ζ еще одно новое значение. В результате поведение системы определится фазовой траекторией, изображенной на фиг. 137, *в*.

Но и это поведение системы, хотя и связанное с наличием точки устойчивого равновесия внутри контура

переключения, не предохраняет систему от ухода изображающей точки за пределы контура, что должно произойти в точке P_2. Таким образом, переключатель должен сработать в третий раз, после чего поведение системы будет определяться фазовыми траекториями, изображенными на фиг. 137, *г*. В этих условиях система становится уже устойчивой и изображающая точка системы из положения P_2 стремится к точке P_3 равновесия. Этот тип действия системы сохранится, так как благодаря устойчивости положения равновесия изображающая точка системы не пересечет контура переключения и, таким образом, не произойдет возбуждения переключающего устройства, призванного изменять параметр ζ.

Поэтому, чтобы заставить систему автоматически искать устойчивый режим, отбрасывая неустойчивые режимы и сохраняя режимы устойчивые, достаточно лишь ввести в систему переключающее устройство и задать определенный *контур переключения*. В остальном последовательность переключений может быть совершенно произвольной. Если устойчивый режим достигается сразу же после первого переключения, то тем лучше для стабилизируемого процесса. Но окончательный результат действия переключателя один и тот же, один ли раз сработал этот переключатель, или три раза. Система всегда достигнет устойчивости. Таким образом, мы оказываемся в состоянии достичь *целеустремленного* поведения системы с помощью чисто механических средств. Такая система автоматически сохраняет устойчивость, и нам нет надобности заботиться об устойчивости при ее проектировании, ибо она сама найдет состояние устойчивости [1]). Таким образом, такая система обладает свойствами, более сильными, чем только свойство устойчивости, и поэтому, следуя Эшби, мы будем называть ее ультраустойчивой.

Наш пример автономной системы с двумя независимыми переменными можно обобщить на случай системы n пере-

---

[1]) Само собой разумеется, что при этом система должна быть устойчивой хотя бы при одном из возможных значений параметра ζ, физически осуществляемых с помощью переключающего устройства. В системах структурно неустойчивых, т. е. неустойчивых при любом ζ, этот прием не приведет к успеху.— *Прим. перев.*

менных y_i, $i=1, \ldots, n$. При этом соответствующую систему дифференциальных уравнений можно записать в виде

$$\frac{dy_i}{dt} = f_i(y_1, y_2, \ldots, y_i, \ldots, y_n; \zeta), \quad i=1, \ldots, n, \qquad (17.3)$$

где ζ — параметр, который принимает различные значения каждый раз, когда фазовая траектория в фазовом пространстве переменных y_i пересекает поверхность переключения. Под поверхностью переключения в данном случае мы понимаем замкнутую гиперповерхность $n-1$ измерений в фазовом пространстве n измерений. Такая система также является ультраустойчивой.

17.2. Пример ультраустойчивой системы. Чтобы продемонстрировать действие ультраустойчивой системы, Эшби построил сравнительно простую систему с четырьмя независимыми переменными, названную им гомеостатической[1]). Четырьмя переменными y_1, y_2, y_3 и y_4 служат углы отклонения четырех магнитов, причем движения последних сильно демпфированы. Положение каждого магнита управляется четырьмя обмотками, каждая из которых питается токами, генерируемыми угловыми положениями всех четырех магнитов.

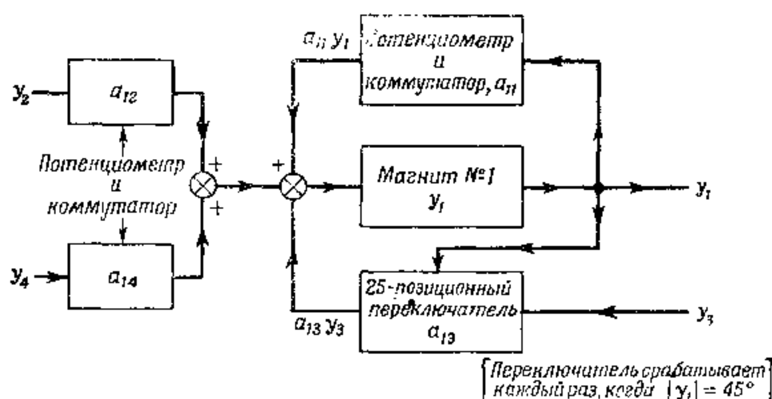
Благодаря сильному демпфированию магниты вращаются медленно, поэтому силами инерции можно пренебречь. Приравняв движущие моменты магнитов, создаваемые обмотками, моменту сил вязкого трения, мы получим уравнения движения магнитов:

$$\left.\begin{aligned}
\frac{dy_1}{dt} &= a_{11}y_1 + a_{12}y_2 + a_{13}y_3 + a_{14}y_4, \\
\frac{dy_2}{dt} &= a_{21}y_1 + a_{22}y_2 + a_{23}y_3 + a_{24}y_4, \\
\frac{dy_3}{dt} &= a_{31}y_1 + a_{32}y_2 + a_{33}y_3 + a_{34}y_4, \\
\frac{dy_4}{dt} &= a_{41}y_1 + a_{42}y_2 + a_{43}y_3 + a_{44}y_4.
\end{aligned}\right\} \qquad (17.4)$$

---

[1]) Ashby W. R., Electronic Eng., **20**, 379 (1948).

В реальной установке абсолютные величины коэффициентов a_{11}, a_{12}, ..., a_{14} можно изменять при эксперименте путем регулирования токов в обмотках с помощью потенциометра, знаки же этих коэффициентов можно изменять посредством коммутаторов, включенных в цепь обмоток. Кроме того, одна из обмоток каждого магнита питается током через посредство переключателя, имеющего 25 возможных положений. При отклонении одного из магнитов



Фиг. 138

на 45° в любую сторону соответствующий переключатель срабатывает и принимает произвольное новое положение. Таким образом, четыре коэффициента среди a_{ij} (i, $j = 1$, 2, 3, 4) могут принимать каждый одно из 25 значений, определяемых произвольным срабатыванием четырех переключателей. Для каждого магнита или, лучше сказать, для каждой переменной y_i мы можем построить блок-схему, подобную приведенной на фиг. 138.

Таким образом, поверхностью переключения гомеостата служит «куб» в четырехмерном пространстве с центром в начале координат, ребра которого соответствуют угловому перемещению на 90°. Каждому сочетанию коэффициентов a_{ij}, задаваемому экспериментатором, соответствуют $25^4 = 390\,625$ сочетаний значений четырех коэффициентов, определяемых положениями четырех

переключателей. Поэтому каждой входной настройке коэффициентов отвечает 390 625 режимов гомеостата, причем некоторые из них будут устойчивыми, а другие — неустойчивыми. Впрочем, неустойчивые режимы система автоматически отбрасывает.

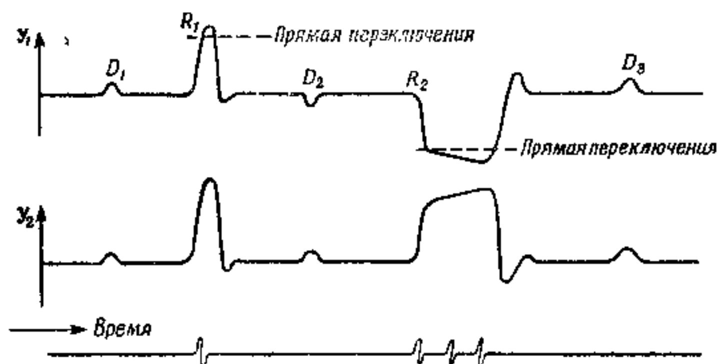
Теперь можно продемонстрировать свойство ультраустойчивости. Сначала для простоты рассматривается только один блок, охваченный обратной связью на себя через посредство единственного переключателя; другие обмотки отключены. Поведение одного магнита, изолированного от остальных, графически изображено на фиг. 139, где верхняя



Фиг. 139

кривая показывает угловое отклонение магнита, а нижняя кривая характеризует действие переключателя. В точке D_1 экспериментатор вручную возмущает движение магнита, но, так как положение переключателя случайно соответствует устойчивой системе, это отклонение быстро выправляется. В точке R_1 экспериментатор меняет полярность цепи обратной связи, в результате чего в старом положении переключателя система оказывается неустойчивой и угловое отклонение магнита достигает линии срабатывания переключателя (прямая переключения показана на фиг. 139 пунктиром). Переключатель срабатывает. После одного скачка ключа переключателя система становится устойчивой, о чем свидетельствует результат пробного возмущения в точке D_2. В точке R_2 экспериментатор вновь меняет полярность цепи обратной связи. На этот раз ключ переключателя поочередно принимает уже четыре случайных положения, прежде чем система становится устойчивой. Об устойчивости состояния системы после четвертого переключения свидетельствует результат пробного возмущения, осуществленного в точке D_3.

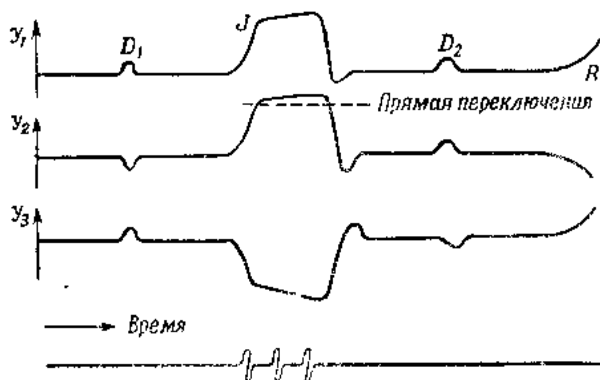
Перейдем к следующему примеру, в котором мы рассматриваем движение двух взаимодействующих магнитов, угловые координаты которых суть y_1 и y_2. Коэффициент a_{21} устанавливается экспериментатором, а коэффициент a_{12} определяется положением переключателя. Значения всех остальных коэффициентов устанавливаются равными нулю.



Фиг. 140

Каждой настройке коэффициента a_{21} отвечают 25 различных режимов, обусловливаемых 25 положениями переключателя. Результаты эксперимента можно графически представить с помощью фиг. 140, где две верхние кривые изображают отклонения переменных y_1 и y_2, а нижняя кривая характеризует действие переключателя. Реакция системы на пробное возмущение, осуществляемое в точке D_1, указывает на то, что данное состояние системы устойчиво, а отклонения y_1 и y_2 имеют одинаковые знаки. В точке R_1 происходит изменение знака коэффициента a_{21}, что достигается изменением полярности обмотки магнита с координатой y_2, питаемой током от магнита с координатой y_1. Это действие приводит к неустойчивости, и в результате отклонение y_1 достигает линии переключения. Положение здесь выправляется сразу после первого срабатывания переключателя. Как показывает реакция системы на пробное возмущение, приложенное в точке D_2, система теперь становиться устойчивой, а отклонения y_1 и y_2 оказываются различными по знаку, что и следовало ожидать. В точке R_2 знак коэффициента a_{21}

опять изменяется и совпадает теперь с его знаком в точке D_1: положение переключателя соответствует при этом неустойчивой системе. Теперь переключатель срабатывает три раза подряд, прежде чем достигается устойчивый режим системы. Как видно из результатов испытания, проведенного в точке D_3, после третьего срабатывания переключателя система действительно становится устойчивой, причем оба отклонения y_1 и y_2 оказываются одного и того же знака.



Фиг. 141

В качестве примера способности ультраустойчивой системы применяться даже к непредвиденным условиям (т. е. к условиям, не рассматривавшимся до физического осуществления машины), исследуем поведение системы при обстоятельствах, характеризуемых графиками на фиг. 141. Здесь рассматриваются движения трех взаимодействующих магнитов. В начале состояние системы устойчиво, о чем свидетельствует реакция системы на возмущение, приложенное в точке D_1. Отклонения y_1 и y_3 имеют одинаковые знаки, а знак отклонения y_2 им противоположен. Пусть в точке J мы подвергли гомеостат неожиданному воздействию, скрепив магниты с координатами y_1 и y_2 таким образом, что начиная с этого момента они должны двигаться как одно тело. Эта дополнительная связь, наложенная на систему, противоречит предыдущей установке переключателя. Система становится неустойчивой и в результате претерпевает большое отклонение, от которого происходит

срабатывание переключателя. Далее система последовательно отбрасывает три неустойчивых режима и только после этого достигает устойчивости, о которой свидетельствует ее реакция на пробное возмущение, приложенное в точке D_2. В точке R жесткая связь между магнитами с координатами y_1 и y_2 разрушается и система вновь становится неустойчивой и требует нового срабатывания переключателя.

17. 3. Вероятность нахождения устойчивого режима. В предыдущих разделах мы описали характерную способность приспособления ультраустойчивой системы, а именно, способность искать устойчивые режимы при любых обстоятельствах. Но при этом возникает вопрос: всегда ли увенчиваются успехом эти поиски устойчивого режима? Какова вероятность успеха? В случае автономной системы n переменных, описываемой системой уравнений (17.3), каждому значению параметра ζ, принимающего дискретные значения, отвечает обычная динамическая система. Следовательно, полному диапазону изменения параметра ζ отвечает некоторое множество автономных динамических систем n переменных. Определим вероятность нахождения устойчивого режима внутри поверхности переключения следующим образом. Возьмем в фазовом пространстве наудачу точку $P(y_i)$ и окружим эту точку бесконечно малой окрестностью dV. Тогда мы можем найти дифференциальную вероятность dp как ту часть множества динамических систем, которая обладает устойчивым равновесием внутри объема dV. Теперь проинтегрируем dp по всему фазовому пространству, ограниченному поверхностью переключения. В результате получим *общую вероятность p нахождения устойчивого режима*, соответствующую заданной поверхности переключения.

Едва ли нужно отмечать, что фактическое вычисление этой общей вероятности представляет собой очень трудную математическую задачу. Чтобы получить некоторое представление о величине этой вероятности, Эшби провел элементарный подсчет для множества линейных систем вида

$$\frac{dy_i}{dt} = a_{ij} y_j, \qquad i, j = 1, 2, \ldots, n. \tag{17.5}$$

В этом случае имеется только одна точка равновесия, а именно, начало координат. Вопрос об устойчивости системы

решается исследованием характеристического уравнения

$$a_{ij} - \delta_{ij} \lambda^i = 0, \qquad \begin{matrix} \delta_{ij} = 1, & i = j, \\ \delta_{ij} = 0, & i \neq j. \end{matrix} \tag{17.6}$$

Если вещественные части всех корней λ отрицательны, то система устойчива. Корень λ называется, как известно, собственным числом матрицы a_{ij}. Таким образом, вероятность получения устойчивого режима равна вероятности того, что собственные числа матрицы a_{ij}, принадлежащей к некоторому множеству, все будут иметь отрицательные действительные части. Эшби рассмотрел простейшее возможное распределение, а именно прямоугольное распределение. В этом случае каждый элемент a_{ij} матрицы может с одинаковой вероятностью принимать значение, равное любому целому числу от -9 до $+9$ включительно. Выбор чисел a_{ij} был произведен с помощью таблицы случайных чисел. При $n = 1$ вероятность нахождения устойчивого режима равна, очевидно, $^1/_2$. Для систем более высоких порядков, Эшби проводил исследование на устойчивость с помощью правила Гурвица; результаты его поисков сведены в табл. 3. Как следует из этой таблицы, вероятность p наступления устойчивости убывает с повышением порядка системы. Приближенно эта вероятность равна $(^1/_2)^n$.

Таблица 3

n	Число испытаний	Число устойчивых случаев	Число устойчивых случаев (в процентах)
2	320	77	24
3	100	12	12
4	100	1	1

Вероятность нахождения устойчивого режима можно повысить, если ограничить возможные значения коэффициентов a_{ij} некоторой областью их благоприятных значений. Пусть, например, все диагональные элементы равны нулю или отрицательным числам. Тогда система всегда будет устойчивой, если только переменные

не взаимодействуют. Вероятность наступления устойчивости в случае одной переменной, т. е. при $n=1$, в данных условиях равна, конечно, единице. При $n=2$ эта вероятность равна $3/4$. Результаты соответствующих испытаний, проведенных Эшби, приведены в табл. 4.

Таблица 4

n	Число испытаний	Число устойчивых случаев	Число устойчивых случаев (в процентах)
2	120	87	72
3	100	55	55

Таким образом, вероятность p нахождения устойчивого режима системы оказывается более высокой, но, тем не менее, убывает с ростом числа переменных.

Если исходить из этих наблюдений, то представляется справедливым мнение о том, что вероятность устойчивости произвольно построенной системы неизменно убывает с ростом сложности системы. Поэтому вероятность того, что сложная система будет неустойчивой, намного более вероятности того, что такая система будет устойчивой.

17.4. Установившиеся поля фазовых траекторий. Следуя Эшби, дадим теперь определенное название режиму поведения системы, отвечающему каждой данной настройке параметра ζ. Совокупность траекторий в фазовом пространстве, характеризующих поведение системы, назовем *полем фазовых траекторий*. Различным значениям параметра отвечают различные поля фазовых траекторий. То устойчивое поле, к которому мы окончательно придем после переключений параметра, назовем *установившимся полем*. Таким образом, действие ультраустойчивой системы главным образом состоит в разыскании установившегося поля. Поэтому исключительно важно знать среднее число N переключений, необходимое для достижения установившегося поля. Это число N определяется очень просто через посредство вероятности p нахождения устойчивого режима ультраустойчивой системы. Очевидно, что

вероятность достижения установившегося поля после первого же срабатывания переключателя равна самой вероятности p, а вероятность получения неустойчивого режима равна $q = 1 - p$. Если переключение происходит идеально случайным образом, то вероятность устойчивости второго поля попрежнему равна p, а вероятность неустойчивости второго поля равна q. Таким образом, относительная вероятность того, что второе поле будет установившимся, равна pq. Относительная же вероятность того, что второе поле не будет установившимся, равна q^2. Отсюда мы приходим к выводу, что относительная вероятность достижения установившегося поля после m-го срабатывания переключателя равна pq^{m-1}. Таким образом, среднее число N срабатываний переключателя для достижения установившегося поля равно

$$N = \frac{\sum_{m=1}^{\infty} mpq^{m-1}}{\sum_{m=1}^{\infty} pq^{m-1}} = \frac{(1-q)^{-2}}{(1-q)^{-1}} = \frac{1}{1-q} = \frac{1}{p}. \tag{17.7}$$

Если p очень мало, то в очень больших системах число N срабатываний переключателя, необходимое для достижения устойчивого режима, окажется очень большим. Это значит, что процесс поисков установившегося поля будет длительным и, повидимому, пойдет по окольному пути.

Установившееся поле может носить особый характер в том смысле, что только небольшая часть фазовых траекторий будет стремиться к точке равновесия, тогда как остальные фазовые траектории будут расходиться от точки равновесия и пересекать поверхность переключения. Такое поле будет обладать свойством установившегося поля только в том случае, если фазовая траектория, ведущая от исходной точки на этой поверхности, окажется принадлежащей к первому небольшому множеству траекторий. Мы покажем здесь, что такие особые установившиеся поля не являются для нас благоприятными. Если среди всех возможных полей ультраустойчивой системы имеется некоторая часть таких особых полей, то число таких полей, пригодных для использования в качестве установившихся, намного меньше. Чтобы показать это, обозначим через k часть поверхности

переключения, обладающую тем свойством, что фазовые траектории, исходящие из точек этой части поверхности, будут достигать точки равновесия.

Например, для полей, изображенных на фиг. 137, *а*, *б*, $k=0$. Для поля, изображенного на фиг. 137, *в*,k приблизительно равно 1/2. Для поля, показанного на фиг. 137, *г*, $k=1$. Пусть теперь $f(k)\,dk$ — часть всех возможных полей ультраустойчивой системы, для которой значение k заключено между k и $k+dk$. Таким образом, $f(k)$ представляет собой функцию распределения возможных полей ультраустойчивой системы и, по определению,

$$\int_0^1 f(k)\,dk = 1. \tag{17.8}$$

Поэтому, так как установившееся поле могут создать только те фазовые траектории, которые выходят из части k поверхности переключения, вероятность получения установившегося поля с величиной k, заключенной в пределах между k и $k+dk$, равна $kf(k)\,dk$. Итак, *функция распределения* $g(k)$ установившихся полей связана с функцией $f(k)$ распределения возможных полей ультраустойчивой системы с помощью соотношения

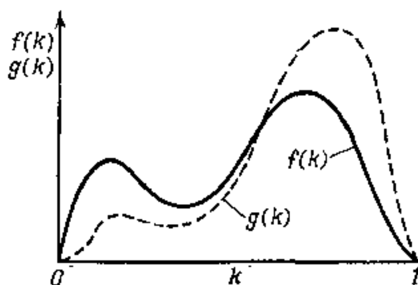
$$g(k) = \frac{kf(k)}{\int_0^1 k'f(k')\,dk'}. \tag{17.9}$$

Очевидно, что

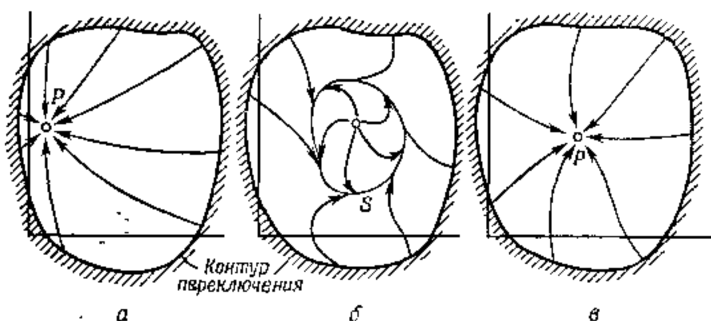
$$\int_0^1 g(k)\,dk = 1. \tag{17.10}$$

Таким образом, установившиеся поля отчетливо сгущаются в сторону бо́льших значений k, что и показывает график на фиг. 142. Следовательно, особые установившиеся поля встречаются редко. Однако истинное распределение установившихся полей не обязательно совпадает с их теоретическим распределением, соответствующим соотношению (17.9). Причина этого явления состоит в том, что на всякое состояние равновесия в установившемся поле дей-

ствуют случайные возмущения. Если точка равновесия в фазовом пространстве лежит очень близко от поверхности переключения, то даже относительно малое возмущение может в некоторый момент вывести систему за поверхность переключения и тем самым разрушить данное установившееся поле. Поэтому, чтобы установившееся поле обладало значительной вероятностью сохраниться под действием случайных возмущений, точка равновесия должна располагаться в центральной части области, ограниченной поверхностью включения. Например, как показано на фиг. 143, поле *в* заведомо более устойчиво по сравнению с полями *а* и *б*. Поле *б* содержит как точку неустойчивого равновесия, так и предельный цикл.



Фиг. 142



Фиг. 143

Чтобы облечь это понятие устойчивости поля по отношению к случайным возмущениям в форму количественных соотношений, введем вероятность σ сохранения установившегося поля после одиночного возмущения. Если поле содержит только одну точку P устойчивого равновесия (фиг. 143) и если задана функция распределения

случайных возмущений состояния равновесия в точке P, — например функция распределения Гаусса, — то σ просто сводится к интегралу от этой функции распределения возмущения, взятому по той части фазового пространства, которая охвачена поверхностью переключения. Если в установившемся поле содержится предельный цикл S, то σ представляет собой среднее значение вероятности сохранения поля; при этом каждая точка на предельном цикле принимается за точку равновесия, а указанное среднее значение вероятности взвешивается в соответствии с временем нахождения системы в соответствующей точке предельного цикла. После этого мы можем каждому установившемуся полю поставить в соответствие вероятность σ. Пусть $\varphi(\sigma)$ есть функция распределения установившихся полей по σ; тогда вероятность нахождения установившегося поля со значением σ в промежутке между σ и $\sigma + d\sigma$ равна $\varphi(\sigma)\,d\sigma$. Таким образом

$$\int_0^1 \varphi(\sigma)\,d\sigma = 1. \tag{17.11}$$

Однако истинная функция распределения установившегося поля равна $\psi(\sigma)$, причем

$$\int_0^1 \psi(\sigma)\,d\sigma = 1. \tag{17.12}$$

Чтобы выразить $\psi(\sigma)$ через посредство $\varphi(\sigma)$, заметим прежде всего, что распределение $\psi(\sigma)$, являющееся истинным распределением установившегося поля, не зависит от случайных возмущений. Далее, обратим внимание на то, что после одного возмущения те поля, для которых вероятность σ заключена между σ и $\sigma + d\sigma$ и относительное число которых составляет $\psi(\sigma)\,d\sigma$, будут обладать вероятностью σ сохранения и вероятностью $1-\sigma$ разрушения. Полное относительное число полей, уничтоженных в результате воздействия со стороны одного случайного возмущения, таким образом, равно

$$\int_0^1 (1-\sigma)\,\psi(\sigma)\,d\sigma.$$

Закон распределения новых установившихся полей отвечает закономерности, справедливой для теоретических установившихся полей, т. е. определяется функцией $\varphi(\sigma)$. Поэтому полное относительное число установившихся полей, сохранившихся после одного случайного возмущения в промежутке от σ до $\sigma + d\sigma$, равно

$$\sigma\psi(\sigma)\,d\sigma + \varphi(\sigma)\,d\sigma \int_0^1 (1-\sigma')\,\psi(\sigma')\,d\sigma'.$$

Но это относительное число пропорционально $\psi(\sigma)\,d\sigma$, так как случайные возмущения не влияют на итоговое распределение $\psi(\sigma)$. Таким образом, обозначая через C постоянный коэффициент пропорциональности, мы можем написать

$$C\left[\sigma\psi(\sigma) + \varphi(\sigma)\int_0^1 (1-\sigma')\,\psi(\sigma')\,d\sigma'\right] = \psi(\sigma).$$

Интегрируя это выражение по σ от $\sigma = 0$ до $\sigma = 1$, можно показать, что при выполнении соотношений (17.11) и (17.12) C равно единице. Поэтому

$$\sigma\psi(\sigma) + \varphi(\sigma)\int_0^1 (1-\sigma')\,\psi(\sigma')\,d\sigma' = \psi(\sigma).$$

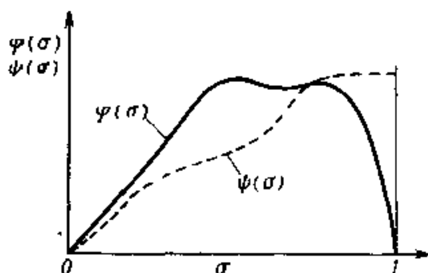
В этом соотношении интеграл представляет собой постоянную, не зависящую от σ. Таким образом, функция $\psi(\sigma)$ пропорциональна $\varphi(\sigma)/(1-\sigma)$, так что, принимая во внимание равенство (17.12), мы будем иметь

$$\psi(\sigma) = \frac{\varphi(\sigma)}{(1-\sigma)\displaystyle\int_0^1 \frac{\varphi(\sigma')\,d\sigma'}{1-\sigma'}}. \tag{17.13}$$

С помощью соотношения (17.13) можно определять истинную функцию распределения установившегося поля по теоретической функции распределения этого поля. Так как теоретическую функцию распределения можно вычислить по функции распределения всех возможных полей

ультраустойчивой системы посредством соотношения (17.9), то мы можем, по крайней мере теоретически, найти истинную функцию $\psi(\sigma)$ распределения установившегося поля по заданным свойствам ультраустойчивой системы.

К тому же, как показывает соотношение (17.13), истинная функция $\psi(\sigma)$ распределения установившегося поля принимает наибольшие значения при больших значениях σ, и это явление здесь выражено гораздо более резко, чем в случае функции $\varphi(\sigma)$, что мы и ожидали, исходя из наших предыдущих интуитивных рассуждений. Это явление графически изображено на фиг. 144. Следует



Фиг. 144

отметить, что от конкретной формы случайных возмущений зависит только функция распределения $\varphi(\sigma)$, тогда как соотношение между функциями $\psi(\sigma)$ и $\varphi(\sigma)$, выраженное равенством (17.13), не зависит от этой формы и справедливо при любом виде функции распределения возмущений.

17.5. Мультиустойчивые системы. Как показано в предыдущем разделе, следует ожидать, что число N срабатываний переключателя, необходимое для достижения установившегося поля, равно $1/p$, где p — общая вероятность устойчивости полей данной ультраустойчивой системы. Так как, по нашей оценке, вероятность p убывает в больших системах до очень малых значений, число N может быть очень большим. Например, если система содержит 100 переменных, то $p \approx 1/2^{100}$ и $N = 2^{100} \approx 10^{30}$. Даже если принять, что переключатель срабатывает 10 раз

за одну секунду, то среднее время достижения установившегося поля составит все-таки $3 \cdot 10^{10}$ столетий. Столь длительное время установления вполне можно считать бесконечно большим, и практически ультраустойчивая система никогда не достигнет установившегося поля. Поэтому в больших системах,—т. е. именно там, где принцип ультраустойчивости имеет важное значение для достижения устойчивого режима,—этот принцип оказывается непрактичным.

Чтобы поправить положение, мы должны увеличить вероятность нахождения устойчивого режима системы. Один из способов увеличения этой вероятности основан на компромиссе: система должна быть спроектирована таким образом, чтобы поля ультраустойчивой системы принадлежали к совокупности полей, устойчивых при ожидаемых эксплоатационных условиях. В этом случае при помощи переключателя пришлось бы вносить лишь ограниченную настройку местного характера. Другими словами, синтез такой системы осуществлялся бы с помощью обычных методов, без применения принципа ультраустойчивости. Ультраустойчивость и связанный с ее осуществлением переключатель вводятся лишь в ожидании возмущений. Например, мы можем разработать автопилот для некоторого самолета, исходя из принципов, изложенных в предыдущих главах. Но у нас возникает опасение, что при сборке автопилота может быть допущена ошибка в установке полярности управляющего сигнала от автопилота на рулевую машинку элеронов: тогда при подаче сигнала на опускание элерона этот сигнал в действительности вызовет его подъем. Если эта ошибка действительно допущена при сборке, то автопилот не будет стабилизировать самолет; больше того, система самолет—автопилот окажется неустойчивой, и движение самолета по крену окажется расходящимся. Однако эту ошибку в сборке можно компенсировать путем введения ультраустойчивости самолета в его движение по крену. Тогда в схеме подачи управляющего сигнала будет предусмотрено автоматическое переключение полярности каждый раз, когда крен самолета выйдет за установленные границы, характеризующие контур переключения. Система, снабженная таким переключателем, становится системой ультраустойчивой, обладающей двумя полями: одно

из них устойчивое, другое—неустойчивое. Для достижения устойчивого режима необходимо выполнить, самое большее, одно переключение, несмотря на то, что состояние системы самолет—автопилот может определяться весьма большим числом переменных. Смысл такого компромисса заключается в том, что мы не обязаны предоставлять случаю все переменные. Мы можем предусмотреть в проекте устойчивое действие системы почти при всех условиях и в сильной мере сузить возможный выбор полей системы, обеспечив меры безопасности только в отношении случайностей, которых мы имеем основание опасаться. Следовательно, такой подход и служит компромиссом между обычными принципами синтеза систем автоматического регулирования и принципом ультраустойчивости.

Когда мы имеем дело с живыми существами, то нельзя заранее знать условия, в которых эти существа могут оказаться, и поэтому по отношению к ним не существует возможности повышения вероятности нахождения устойчивости путем сужения выбора полей, характеризующих поведение системы. Эшби открыл другой путь повышения этой вероятности. Он заметил, что в очень сложных системах, насчитывающих большое число переменных, любое отдельное возмущение или изменение режима работы системы непосредственно отражается только на относительно небольшом числе этих переменных. Таким образом, для этих конкретных видов возмущений открывается возможность значительного повышения вероятности нахождения устойчивого режима путем изоляции переменных, непосредственно претерпевших возмущение, от остальных переменных системы и образования из этих изолированных переменных ультраустойчивой системы. Например, если под действием возмущения непосредственно нарушается течение пяти переменных вместо 100, как мы это полагали в начале настоящего раздела, то ожидаемое время нахождения установившегося поля системы составит всего 3,2 сек. при тех же 10 срабатываниях переключателя в секунду. Таким образом, если удастся разбить 100 переменных системы на 20 классов по пяти переменных в каждом классе, с образованием из них 20 раздельных ультраустойчивых систем, то полное время приспособления системы к совершенно новой совокупности рабочих режимов составит $20 \times 3{,}2 = 64$ сек. Этим будет до-

стигнуто огромное сокращение времени поисков устойчивого режима по сравнению с временем в 3.10^{19} столетий, требуемым для этих поисков в случае ультраустойчивой системы со всеми действующими соединениями.

Конечно, система, образованная из 20 отдельных систем, каждая из которых состоит из пяти переменных, не обладает гибкостью и богатством переходных процессов, присущими системе из 100 взаимодействующих связанных переменных. Однако, если любое возмущение действует только на пять переменных, то совокупность этих раздельных систем по пяти переменных можно сделать равносильной системе из 100 переменных, если каждый раз группировать переменные по всем этим совокупностям в соответствии с действием возмущения. Так, если возмущенные переменные суть y_1, y_2, y_3, y_4 и y_5, то эти пять переменных следует сгруппировать в одну ультраустойчивую систему пяти переменных. Если же последующее возмущение отражается на переменных y_2, y_5, y_{10}, y_{98} и y_{99}, то ультраустойчивая система образуется именно из этих пяти переменных. Это явление последовательного перегруппирования переменных в «подсистемы» (совокупности) в соответствии с условиями работы системы Эшби назвал *дисперсией режимов*. Физически дисперсию можно осуществить путем обращения функций $f_i(y_1, y_2, \ldots, y_n; \zeta)$, входящих в уравнение (17.3), в нуль при расположении точки $(y_1, y_2, \ldots, y_n)$ внутри некоторой области в фазовом пространстве. Тогда эти величины y_i будут сохранять постоянное значение во времени, а по отношению к другим переменным они будут играть лишь роль пробных параметров. В нашем примере, рассмотренном выше, при первом возмущении точка $(y_1, y_2, \ldots, y_n)$ расположена в такой области фазового пространства, в которой $f_i=0$ при всех значениях i, кроме значений $i=1$, 2, 3, 4 и 5. При втором возмущении f_i исчезает при всех i, кроме $i=2,5,10,98$ и 99. Такое поведение функций означает, очевидно, лишь то, что для производных dy_i/dt существуют различные пороговые значения. Конечно, такие пороговые значения существуют, как можно ожидать, во всякой физической системе. Поэтому физически дисперсия осуществляется довольно легко.

Систему ультраустойчивых подсистем, образованную с обеспечением возможности дисперсии, Эшби назвал

мультиустойчивой системой. Ясно, что мультиустойчивая система обладает способностью к приспособлению, присущей ультраустойчивой системе, по той простой причине, что она образована из ультраустойчивых подсистем. Однако она отличается от отдельной ультраустойчивой системы с равным числом переменных по продолжительности поисков, требующихся для достижения установившегося поля. Мультиустойчивая система обладает гораздо меньшим временем установления, благодаря чему принцип ультраустойчивости оказывается практически осуществимым. Кроме того, мультиустойчивая система, реагирующая на последовательные возмущения последовательными пробными процессами приспособления, тем самым срабатывает по этапам, т. е. допускает последовательное приспособление. Такое поведение повсюду обнаруживается у живых существ. Далее отметим, что, так как второй и последующие процессы приспособления системы к некоторому возмущению непременно приведут к изменению параметра системы, при повторном возникновении возмущения, тождественного первому, в общем случае процесс в системе уже не будет сводиться к воспроизведению процесса приспособления системы, протекавшего в системе в ходе первого приспособления. В этом, по существу, и состоит дисперсия режимов. Другими словами, по мере того, как система становится «старше», она становится и «мудрее», и ее поведение уже не ограничивается лишь воспроизведением процесса, последовавшего за отдельным преобладающим возмущением, а способно отвечать более широким требованиям.

Глава XVIII

ПРЕДУПРЕЖДЕНИЕ ОШИБОК В СИСТЕМАХ

В предыдущей главе мы изложили методику применения принципа ультраустойчивости для придания системе регулирования свойства нечувствительности по отношению к случайным ошибкам и непредвиденным выходам из строя отдельных элементов системы, что связано с введением простого устройства для изменения характеристик системы при всяком возникновении неустойчивости. Так как ультраустойчивая система будет автоматически разыскивать устойчивый режим, система регулирования, разработанная на этом принципе, в действительности будет содержать как неустойчивые фазовые поля, так и устойчивые фазовые поля. Другими словами, при синтезе ультраустойчивой системы мы не предпринимаем какой-либо попытки провести различие между устойчивостью и неустойчивостью и отделить пригодные фазовые поля от непригодных. При этом ошибки в отработке системы рассматриваются лишь с точки зрения вероятности их возникновения; в других отношениях они не характеризуются никак. В настоящей главе мы подойдем к вопросу о надежности сложной системы регулирования с другой точки зрения: мы преднамеренно введем в систему ошибки и поставим вопрос о том, как следует запроектировать систему, чтобы она давала удовлетворительную отработку, несмотря на ошибки. Иначе говоря, мы желаем знать способ предупреждения ошибок.

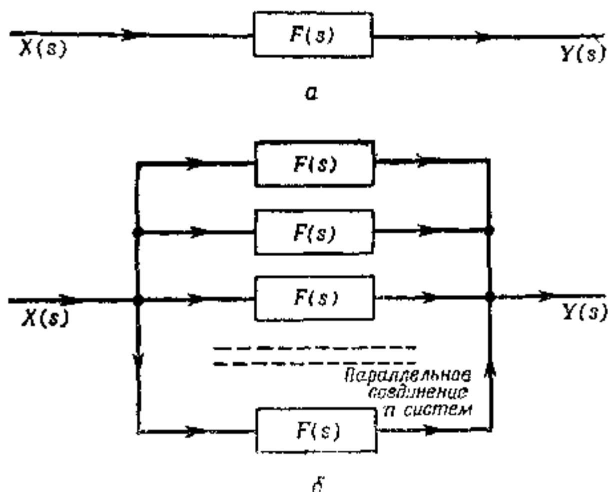
Вопрос о предупреждении ошибок находится в настоящее время в начальной стадии своего развития. Задачу о предупреждении ошибок мы пока можем рассматривать только по отношению к наиболее простым операциям; существующая теория полностью принадлежит Дж. Нейману[1]), и наше рас-

---

[1]) Neumann J., Probabilistic Logics and the Synthesis of Reliable Organisms from Unreliable Components, Printed notes of lectures given at the California Institute of Technology, Pasadena, California, 1952.

суждение в данной главе представляет собой изложение работы Неймана. Наше изложение имеет целью ввести читателя в эту весьма важную область и отметить необходимость дальнейших обширных исследований.

18.1. Обеспечение надежности путем дублирования. Известно, что в общем случае можно повысить надежность системы с помощью простого приема дублирования.



Фиг. 145

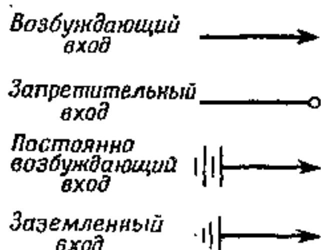
Например, если простая система, схематически изображенная на фиг. 145, *а*, обладает тем свойством, что при отказе в срабатывании эта система не создает на выходе никакого сигнала, то в целях предохранения против неисправности мы можем снабдить систему запасными системами, тождественными с нею; все такие системы соединены параллельно, по схеме, показанной на фиг. 145, *б*. Пусть p — вероятность отказа в срабатывании исходной системы (p — число, заключенное между нулем и единицей); тогда вероятность выхода из строя каждой отдельной системы в параллельном дублирующем соединении также равна p. Если звенья параллельной цепи независимы между собой, то вся цепь выйдет из строя только в том случае, когда из

строя выйдет каждое звено в отдельности. Таким образом, вероятность отказа в срабатывании всей параллельной системы равна p^n. Увеличивая число n систем предохранения, мы можем достичь очень малой величины этой вероятности.

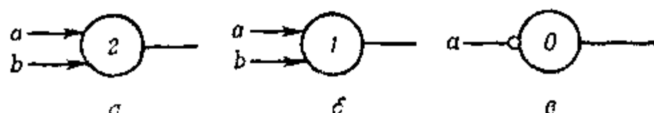
Однако в общем случае неисправность одного из элементов системы регулирования не приводит к исчезновению выходного сигнала. На самом деле влияние такой неисправности сказывается гораздо более вредным образом: система все же создаст некоторый выход, но этот выход будет ложным. В этом случае простое дублирование системы, описанное в предыдущем абзаце, будет недейственным, так как сочетание ложного выхода с правильным выходом доставляет ложный выход. Отсюда следует, что в противоположность случаю отказа при таком неправильном срабатывании системы вероятность неправильного срабатывания параллельной цепи систем равна вероятности неправильного срабатывания одинарной системы и при этом не получается никакого повышения надежности. Поэтому задача о предупреждении ошибок является более трудной и более глубокой, чем это может показаться с первого взгляда. Тем не менее, как мы увидим ниже, принцип дублирования, т. е. необходимость увеличения числа элементов, остается основным принципом. Задача же заключается в открытии новой группировки элементов для предупреждения ошибки, так как простое параллельное соединение, показанное на фиг. 145, *б*, не является действенным.

18.2. Основные элементы. Для упрощения исследования мы не будем рассматривать случай непрерывных входа и выхода и выберем элементарное звено, вход и выход которого могут принимать лишь два дискретных значения, 0 и 1. Это означает, что как вход, так и выход или включен (возбужден), или выключен. Тогда характеристика рассматриваемого звена определяется соотношением между состоянием входа и состоянием выхода. Будем всегда считать, что звено обладает одним выходом, но входов у него может быть несколько. Между выходом и входами имеется также некоторое запаздывание во времени в том смысле, что выходной сигнал возбуждается лишь по истечении некоторого определенного промежутка времени после возбуждения входов. Такое звено обладает свойствами релейного элемента.

Для описания действия этого звена введем четыре вида входов: возбуждающий вход, запретительный вход, постоянно возбуждающий вход и вход, который никогда не возбуждается (заземленный вход). Эти входы графически изображаются символами, приведенными на фиг. 146. Само же звено изображается кружком, по левую сторону которого расположены символы входов, а по правую—его единственный выход. В кружке помещена цифра k, означающая, что для возбуждения выхода число фактически действующих возбуждающих входов должно превышать число фактически действующих запретительных входов на k или больше, чем на k. Так, на

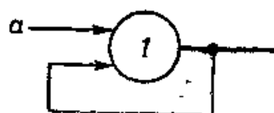


Фиг. 146



Фиг. 147

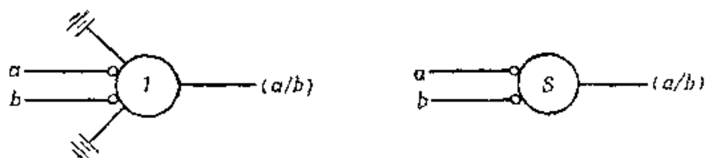
фиг. 147, *а* изображено звено, выход которого включается только при включении обоих входов a и b; такое звено можно назвать ab-звеном. На фиг. 147, *б* показано звено, выход которого включается при включении или входа a, или входа b, или обоих вместе; его можно назвать $a+b$-звеном. На фиг. 147, *в* приведено звено, выход которого включается только при отключении входа a; его можно назвать a^{-1}-звеном. Кстати, если рассматривать входы в качестве условий, при которых некоторое утверждение справедливо (включено) или несправедливо (выключено), то три звена, изображенных на фиг. 147, определяют три основных операции булевской алгебры [1]). В цифровой счетной машине, действующей



Фиг. 148

[1]) Эти устройства иногда называют элементарными логическими блоками.—*Прим. ред.*

по двоичной системе, эти звенья суть основные операционные звенья. Если вычислительные операции необходимо снабдить памятью, то устройство памяти можно выполнить на основе второго звена на фиг. 147, охватив его обратной связью по схеме, изображенной на фиг. 148. Как только вход a возбужден, то возбудится и выход—даже если впоследствии этот вход и будет выключен.



Соединение Шеффера

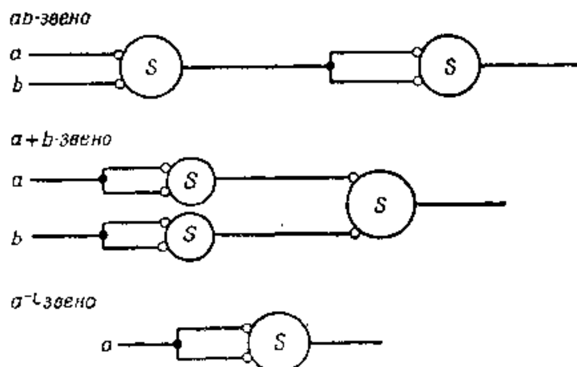
Фиг. 149.

В дальнейших рассуждениях будет невыгодно строить теорию на этих трех основных элементах. Однако эти три звена представляют собой частные случаи одного фундаментального звена. Рассмотрим соединение, показанное на фиг. 149 и называемое соединением или звеном Шеффера. Так как два постоянно возбуждающих входа всегда включены, то мы можем исключить их из схемы и представить это соединение в форме, показанной на этой фигуре справа.

Таким образом, выход звена Шеффера оказывается возбужденным в трех случаях: когда или не включены ни a, ни b, или включен один из входов a или b; но если включены оба входа—a и b, то выход не включен. Три основных звена, рассмотренных в предыдущем абзаце, можно построить с помощью звена Шеффера по способу, изображенному на фиг. 150. Конечно, как ab-звено, так и $a+b$-звено можно представить посредством последовательного соединения двух звеньев Шеффера; таким образом, запаздывание по времени вдвое превысит соответствующее время для a^{-1}-звена, образованного с помощью одного соединения Шеффера. Но так как рассматриваемое здесь понятие запаздывания означает всего лишь то обстоятельство, что выходы предшествуют входам, то точная величина этого запаздывания не имеет значения, и потому операции, различающиеся лишь величиной запаздывания, мы будем считать равносильными. Поэтому

все три основные операции можно изобразить посредством одного звена, образованного с помощью соединения Шеффера.

Такой выбор единого фундаментального звена не является единственным. Помимо соединения Шеффера, возможны и другие основные соединения. Однако соединение, выбранное нами за фундаментальное, оказывается удобным для дальнейшего рассуждения. Мы рассмотрим, прежде всего,

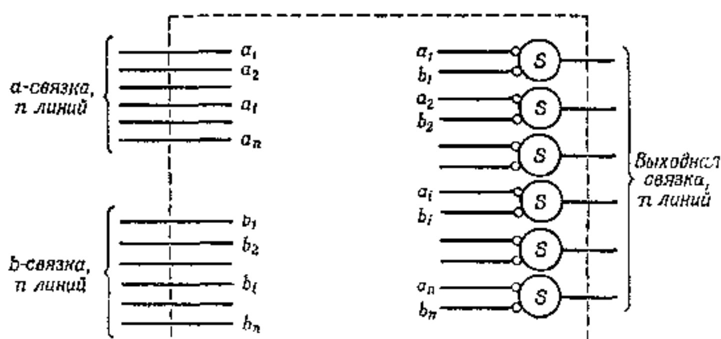


Фиг. 150

предупреждение ошибок в операции, представимой звеном Шеффера; в результате откроется возможность синтеза любой сложной операции, построенной из звеньев Шеффера.

18.3. Метод сложных сочетаний. Обращаясь к принципу дублирования в целях повышения надежности системы, заменим каждый отдельный вход *связкой входов* с n индивидуальными линиями. Таким образом, в системе, представимой с помощью только одного соединения Шеффера, a-входу соответствует n линий, которые обозначены символами a_i $(i = 1, 2, \ldots, n)$; b-входу соответствуют n линий b_i $(i = 1, 2, \ldots, n)$, образующих *выходную связку*. Далее, определим дробь δ, где $0 < \delta < \frac{1}{2}$, так, что если включены или выключены $(1 - \delta)n$ линий выходной связки, то выход рассматривается как включенный или выключенный целиком. Если включены или выключены δn линий выходной связки, то выход рассматривается

соответственно как выключенный или включенный целиком. Всякое промежуточное значение считается результатом неправильного действия системы. Таким образом, δ характеризует критический уровень степени надежности. Существо задачи заключается в построении системы из



Фиг. 151

звеньев Шеффера таким образом, чтобы вероятность неправильного действия ее можно было уменьшить при заданной вероятности ошибок во входных связках и при заданной вероятности неправильного срабатывания отдельных звеньев Шеффера.

В качестве первого приближения к задаче выберем одну линию a_i из входной a-связки и одну линию b_i из входной b-связки и будем считать их входами звена Шеффера. Строение такой системы показано на фиг. 151. Очевидно, что если включены почти все линии обеих входных связок, то почти все линии выходной связки окажутся выключенными. Если почти все линии обеих входных связок выключены, то почти все линии выходной связки окажутся включенными. Такой общий характер действия полной системы представляется удовлетворительным. Однако при более внимательном исследовании оказывается, что это не так.

В самом деле, для выключения выхода соединения Шеффера необходимо возбудить оба входа, поэтому одной ошибки в a-связке или в b-связке будет достаточно для

возникновения ошибки в выходной связке. Следовательно, в предположении, что выход не возбужден, ошибка выхода представляет собой *сумму* ошибок во входных связках. Подобным же образом в предположении, что выход возбужден, а выключена только одна входная связка, ошибка выхода совпадает с ошибкой во входной связке. В предположении, что выход включен, а обе входные связки выключены, для возникновения ошибки в выходной связке требуется *одновременное* возникновение ошибки в обеих входных связках, и поэтому уровень ошибки на входе уже не сохраняется на выходе. Итак, в некоторых условиях наблюдается увеличение ошибки, а в других – ее уменьшение. Такое обстоятельство нежелательно, так как дисперсия ошибок приводит к блужданию числа возбужденных линий выходной связки в неопределенной области между δn линиями и $(1-\delta)n$ линиями, что приводит к повышению вероятности неправильной работы системы.

Для подавления дисперсии ошибок мы можем ввести *восстанавливающее устройство* системы с помощью следующего приема: возьмем каждую линию выходной связки в *исполнительном устройстве*, приведенном на фиг. 151 (исполнительном в том смысле, что агрегат физически осуществляет операцию по схеме Шеффера во всей системе), и расщепим ее на две линии. Таким способом мы получим $2n$ линий. Затем перемешаем эти $2n$ линий так, чтобы порядок следования их стал случайным. Выбирая последовательные пары линий и используя их в качестве входов звена Шеффера, мы опять получим связку n входных линий, что схематически показано на фиг. 152. Пусть в исходной связке содержится $\alpha_0 n$ возбужденных линий. Очевидно, что относительное число не возбужденных выходных линий составляет $\alpha_0\alpha_0=\alpha_0^2$. Тогда относительное число α_1 возбужденных линий равно

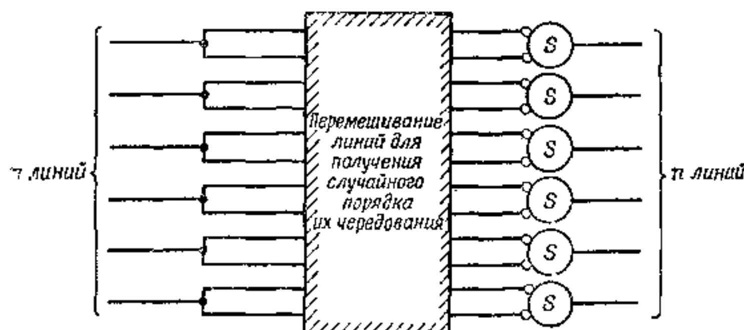
$$\alpha_1=1-\alpha_0^2. \tag{18.1}$$

Обозначим через α_0 вероятность возбуждения исходных линий. Тогда при условии, что n достаточно велико, число α_1 определит вероятность возбуждения перемешанных линий. Но пока мы еще не получим восстанавливающего устройства. Однако, если соединить два таких

устройства последовательно, то вероятность α_2 возбуждения линий на выходе такого соединения выразится равенством

$$\alpha_2 = 1 - \alpha_1^2 = 1 - (1 - \alpha_0^2)^2 = 2\alpha_0^2 - \alpha_0^4. \qquad (18.2)$$

Эта система последовательного соединения уже будет восстанавливающим устройством, о чем свидетельствует следующее рассуждение. На фиг. 153 графически изображена



Фиг. 152

зависимость между вероятностями α_2 и α_0. α_2 равно α_0 при выполнении равенства

$$\alpha_0^4 - 2\alpha_0^2 + \alpha_0 = 0,$$

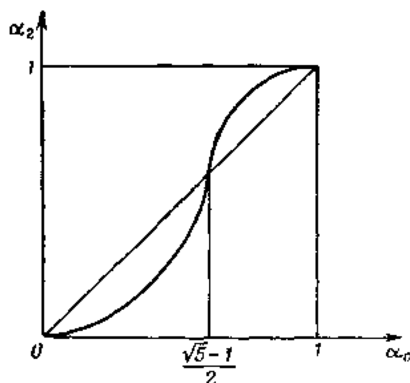
т. е. когда $\alpha_0 = 0$, $\frac{1}{2}(\sqrt{5} - 1)$ или 1. Следовательно, если вероятность α_0 заключена между 0 и

$$\frac{1}{2}(\sqrt{5} - 1) = 0{,}618034,$$

то α_2 меньше, чем α_0; если же α_0 лежит между $\frac{1}{2}(\sqrt{5} - 1)$ и 1, то α_2 больше, чем α_0. Поэтому действие восстанавливающего устройства заключается в сдвиге вероятности возбуждения выхода в сторону граничных значений 0 или 1 и тем самым в уменьшении дисперсии ошибки, вносимой исполнительным устройством.

Таким образом, как следует из предшествующих рассуждений, наша система предупреждения ошибок состоит

из исполнительного устройства, состоящего из n отдельных элементов Шеффера, к которому подключено восстанавливающее устройство, образованное с помощью двух последовательно соединенных агрегатов, выполненных по схеме, показанной на фиг. 152; каждый из этих агрегатов состоит из n элементов Шеффера и приспособления для перемешивания линий в случайном порядке. Поэтому для каждого звена Шеффера, работающего с идеальной точностью, необходимо развернуть данную систему в сложную систему, содержащую $3n$ элементов Шеффера. Мы видим, что при всяком заданном критическом уровне δ и при заданных вероятностях возникновения ошибок во входных линиях и звеньях Шеффера мы можем насколько угодно повысить надежность всей полной системы путем увеличения n. Следовательно, в своей основе принцип повышения надежности попрежнему заключается в дублировании. Однако, как вытекает из нашего исследования в этом пункте, мы уже располагаем определенной методикой построения цепи из этих звеньев. Этот конкретный метод синтеза надежной в срабатывании системы из ненадежных элементов называется методом сложных сочетаний Неймана.



Фиг. 153

18.4. Ошибка в исполнительном устройстве. Подсчитаем теперь ошибку системы, построенной с помощью звеньев Шеффера по методу сложных сочетаний, описанному в предыдущем пункте, и посмотрим, действительно ли осуществляется в этой системе предупреждение ошибок. Заметим прежде всего, что непосредственным источником ошибок служит само отдельно взятое звено Шеффера как в исполнительном устройстве, так и в восстанавливающем устройстве. Обозначим через ε вероятность воз-

никновения ошибки в срабатывании каждого отдельного звена. В предположении, что имеется r параллельно соединенных звеньев Шеффера и что эти звенья срабатывают независимо одно от другого, вероятность неправильного срабатывания каждого из звеньев в этой системе попрежнему остается равной ε. Поэтому в цепи, состоящей из r параллельно соединенных звеньев, в среднем εr звеньев будут срабатывать с ошибкой. Это же число представляет собой и наиболее вероятное число ошибок в цепи из n параллельно соединенных звеньев.

Вероятность появления какого-либо другого числа ошибок будет уже меньше. В самом деле, задача об определении вероятности $p_0(\rho, \varepsilon, r)$ возникновения ρ ошибок в цепи из r параллельно соединенных элементов при заданной вероятности ε выхода из строя каждого звена в отдельности представляет собой классическую задачу о случайном выборе. Как известно[1]), при больших r

$$p_0(\rho, \varepsilon, r) \approx \frac{1}{\sqrt{2\pi}\sqrt{\varepsilon(1-\varepsilon)r}} e^{-\frac{1}{2}\frac{(\rho-\varepsilon r)^2}{\varepsilon(1-\varepsilon)r}}. \tag{18.3}$$

Поэтому распределение вероятности $p_0(\rho, \varepsilon, r)$ представляет собой нормальное распределение со средним значением εr и средним отклонением $\sqrt{\varepsilon(1-\varepsilon)r}$.

Другим источником ошибок в исполнительном устройстве служат неправильности в подключении линий от входных связок к отдельным звеньям Шеффера. Пусть, например, возбуждено относительное число ξ линий входной a-связки (фиг. 151) и относительное число η линий входной b-связки; тогда следует ожидать, что аналогичное число линий выходной связки будет запрещено.

Но если некоторая линия a_i в i-м звене возбуждена, а линия b_i окажется невозбужденной, то выход i-го звена попрежнему останется возбужденным — даже в отсутствие ошибки в самом звене. Пусть ζ — относительное число возбужденных линий выходной связки исполнительного устройства. Тогда число запрещенных выходных линий равно $(1-\zeta)n$. Полное число возбужденных линий во вход-

[1]) См., например, Margenau H., Murphy G. M., The Mathematics of Physics and Chemistry, New York, 1943, p. 422.

ной a-связке равно ξn, а во входной b-связке равно ηn. Однако число действующих, т. е. правильно подключенных входных a-линий, составляет только $(1-\zeta)n$. Разность $[\xi-(1-\zeta)]n$ представляет собой число недейственных линий. Число недейственных входных b-линий составляет $[\eta-(1-\zeta)]n$.

Поэтому $(1-\zeta)n$ выходных линий поключены правильно, $[\xi-(1-\zeta)]n$ выходных линий не действенны из-за неправильного подключения возбужденных линий a-входа, $[\eta-(1-\zeta)]n$ выходных линий не действенны из-за неправильного подключения возбужденных линий b-входа; наконец, число выходных линий, не действенных из-за наличия во входах a и b невозбужденных входных линий, равно остатку

$$\{1-(1-\zeta)-[\xi-(1-\zeta)]-[\eta-(1-\zeta)]\}n=(2-\xi-\eta-\zeta)n.$$

Таким образом, число возможных действенных сочетаний при такой классификации выхода равно[1])

$$\frac{n!}{[(1-\zeta)n]!\,\{[\xi-(1-\zeta)]n\}!\,\{[\eta-(1-\zeta)]n\}!\,[(2-\xi-\eta-\zeta)n]!}.$$

Со стороны входа число возможных сочетаний n входных a-линий, из которых ξn возбуждены, а $(1-\xi)n$ не возбуждены, составляет

$$\frac{n!}{(\xi n)!\,[(1-\xi)n]!}.$$

Число возможных сочетаний n входных b-линий при ηn возбужденных линиях и $(1-\eta)n$ не возбужденных линиях равно:

$$\frac{n!}{(\eta n)!\,[(1-\eta)n]!}.$$

Поэтому при относительных числах ξ и η возбужденных входов отдельных звеньев Шеффера в условиях их идеальной работы вероятность p_1 возбуждения относительного

[1]) См., например, стр. 415 только что цит. работы

числа ζ линий на выходе равна

$$p_1(\xi, \eta, \zeta; n) =$$

$$= \frac{\dfrac{n!}{[(1-\zeta)n]!\,\{[\xi-(1-\zeta)]\,n\}!\,\{[\eta-(1-\zeta)]\,n\}!\,[(2-\xi-\eta-\zeta)\,n]!}}{\dfrac{n!}{(\xi n)!\,[(1-\xi)\,n]!}\cdot\dfrac{n!}{(\eta n)!\,[(1-\eta)\,n]!}} =$$

$$= \frac{(\xi n)!\,[(1-\xi)\,n]!\,(\eta n)!\,[(1-\eta)\,n]!}{[(1-\zeta)\,n]!\,\{[\xi-(1-\zeta)]\,n\}!\,\{[\eta-(1-\zeta)]\,n\}!\,[(2-\xi-\eta-\zeta)\,n]!\,n!}. \qquad (18.4)$$

Очевидно, что для того, чтобы это выражение имело смысл, ни один из четырех классов выходных линий, рассмотренных несколько выше в этом пункте, не должен содержать число линий, меньшее нуля. Каждый раз, когда это условие не выполняется, вероятность p_1 падает до нуля. Это означает, что p_1 равно нулю каждый раз, когда нарушаются следующие условия:

$$\left.\begin{aligned} 1-\zeta &> 0, \\ \xi-(1-\zeta) &> 0, \\ \eta-(1-\zeta) &> 0 \\ \text{и} \qquad 2-\xi-\eta-\zeta &> 0. \end{aligned}\right\} \qquad (18.5)$$

Приступим теперь к упрощению выражений в правых частях равенств (18.4) в предположении, что n — достаточно большое число. При больших n факториал можно приближенно оценить его асимптотическим выражением по формуле Стирлинга:

$$n! \approx \sqrt{2\pi}\, e^{-n} n^{n+\frac{1}{2}}. \qquad (18.6)$$

С помощью равенства (18.6) запишем p_1 приближенно в виде

$$p_1(\xi, \eta, \zeta; n) = \frac{1}{\sqrt{2\pi n}} \sqrt{a}\, e^{-\theta n}, \qquad (18.7)$$

где

$$a = \frac{\xi(1-\xi)\,\eta\,(1-\eta)}{(\xi+\zeta-1)(\eta+\zeta-1)(1-\zeta)(2-\xi-\eta-\zeta)} \qquad (18.8)$$

и

$$\theta=(\xi+\zeta-1)\ln(\xi+\zeta-1)+(\eta+\zeta-1)\ln(\eta+\zeta-1)+\\+(1-\zeta)\ln(1-\zeta)+(2-\xi-\eta-\zeta)\ln(2-\xi-\eta-\zeta)-\\-\xi\ln\xi-(1-\xi)\ln(1-\xi)-\eta\ln\eta-(1-\eta)\ln(1-\eta). \quad (18.9)$$

Дифференцируя θ по ζ, получаем

$$\frac{\partial\theta}{\partial\zeta}=\ln\frac{(\xi+\zeta-1)(\eta+\zeta-1)}{(1-\zeta)(2-\xi-\eta-\zeta)} \quad (18.10)$$

и

$$\frac{\partial^2\theta}{\partial\zeta^2}=\frac{1}{\xi+\zeta-1}+\frac{1}{\eta+\zeta-1}+\frac{1}{1-\zeta}+\frac{1}{2-\xi-\eta-\zeta}. \quad (18.11)$$

С помощью этих равенств мы находим, что при $\zeta=1-\xi\eta$ $\theta=\dfrac{\partial\theta}{\partial\zeta}=0$. Кроме того, в силу условий (18.5), производная $\partial^2\theta/\partial\zeta^2$ всегда положительна. Поэтому функция θ имеет единственный нуль, который находится в точке $\zeta=1-\xi\eta$. Тогда при очень больших n мы должны рассматривать функцию θ только вблизи ее нуля, что обусловливается экспонентой с отрицательным показателем в правой части (18.7). Но в нуле функции θ, где $\zeta=1-\xi\eta$,

$$\frac{\partial^2\theta}{\partial\zeta^2}=\frac{1}{\xi(1-\eta)}+\frac{1}{\eta(1-\xi)}+\frac{1}{\xi\eta}+\frac{1}{(1-\xi)(1-\eta)}=\frac{1}{\xi(1-\xi)\eta(1-\eta)}.$$

Таким образом, вблизи значения $\zeta=1-\xi\eta$ приближением функции θ служит выражение

$$\theta\sim\frac{1}{2}\frac{[\zeta-(1-\xi\eta)]^2}{\xi(1-\xi)\eta(1-\eta)}. \quad (18.12)$$

При большом n величина a представляет собой функцию от ζ, изменяющуюся медленно по сравнению с экспонентой в правой части равенства (18.7). Таким образом, мы можем взять значение a в точке $\zeta=1-\xi\eta$, т. е.

$$a\sim\frac{1}{\xi(1-\xi)\eta(1-\eta)}. \quad (18.13)$$

Поэтому в конечном счете приближенное выражение вероятности $p_1(\xi,\eta,\zeta;n)$ при очень больших n принимает

вид

$$p_1(\xi, \eta, \zeta; n) \approx \frac{1}{\sqrt{2\pi\xi(1-\xi)\eta(1-\eta)n}} e^{-\frac{1}{2}\frac{[\zeta-(1-\xi\eta)]^2 n}{\xi(1-\xi)\eta(1-\eta)}}. \quad (18.14)$$

Отсюда следует, что p_1 также представляет собой нормальное распределение по отношению к ζ. Среднее значение достигается в точке $(1-\xi\eta)n$, а среднее отклонение равно $\sqrt{\xi(1-\xi)\eta(1-\eta)n}$.

Преобразуем выражение вероятности $p_1(\xi, \eta, \zeta; n)$ к окончательному виду путем перехода к непрерывной функции распределения $W(\zeta; \xi, \eta; n)$ при больших значениях n. Если $W(\zeta; \xi, \eta; n)\,d\zeta$ представляет собой вероятность возбуждения числа выходных линий, заключенных между ζn и $\zeta n + 1 = n(\zeta + 1/n)$, то $d\zeta = 1/n$, и эта вероятность в точности равна $p_1(\xi, \eta, \zeta; n)$. Поэтому

$$W(\zeta; \xi, \eta; n) = np_1 = \frac{1}{\sqrt{2\pi\xi(1-\xi)\eta(1-\eta)/n}} e^{-\frac{1}{2}\left[\frac{\zeta-(1-\xi\eta)}{\sqrt{\xi(1-\xi)\eta(1-\eta)/n}}\right]^2}. \quad (18.15)$$

Следовательно, $W(\zeta; \xi, \eta; n)$ в общем случае выражает вероятность нахождения относительного числа возбужденных линий на выходе между ζ и $\zeta + d\zeta$ при заданных относительных числах ξ и η возбужденных входных связок. Объем этих связок определяется числом n линий. W представляет собой гауссово распределение со средним значением $1-\xi\eta$ и средним отклонением $\sqrt{\xi(1-\xi)\eta(1-\eta)/n}$. Этот результат можно записать и в другой равносильной форме:

$$\zeta = (1-\xi\eta) + \sqrt{\frac{\xi(1-\xi)\eta(1-\eta)}{n}}\,y, \quad (18.16)$$

где y обозначает случайную переменную с нормальным гауссовым законом распределения, причем среднее значение равно нулю, а среднее отклонение — единице. Тогда соотношение (18.16) означает, что функция ζ обладает гауссовым законом распределения со средним значением $1-\xi\eta$ и средним отклонением $\sqrt{\xi(1-\xi)\eta(1-\eta)/n}$. Соотношения (18.15) и (18.16) выражают это же самое обстоя-

тельство. Однако соотношение (18.16) более удобно в приложениях.

Теперь мы можем учесть оба источника ошибок совместно и включить в распределение функции ζ, описываемое соотношением (18.16), влияние несовершенства отдельных звеньев Шеффера. Перепишем приближенное равенство (18.3) в виде, аналогичном равенству (18.16)

$$\rho = \varepsilon r + \sqrt{\varepsilon(1-\varepsilon) r}\, y. \qquad (18.17)$$

Вспомним теперь, что наше исполнительное устройство содержит два класса звеньев Шеффера. Выход одного из видов этих звеньев, по предположению, возбужден; число таких звеньев равно $\zeta\eta$. Возникновение одной ошибки уменьшает число возбужденных линий на единицу.

В силу равенства (18.17) q ошибок в звеньях этого вида определяются соотношением

$$q = \varepsilon\zeta n + \sqrt{\varepsilon(1-\varepsilon)\zeta n}\, y. \qquad (18.18)$$

Звенья другого вида в числе $(1-\zeta)n$ предполагаются невозбужденными. При возникновении одной ошибки число возбужденных линий увеличится на единицу. q' ошибок в звеньях этого вида распределены по закону

$$q' = \varepsilon(1-\zeta)n + \sqrt{\varepsilon(1-\varepsilon)(1-\zeta)n}\, y. \qquad (18.19)$$

Следовательно, разность $q'-q$ представляет собой дополнительное число возбужденных линий на выходе, обусловленное ошибками в самих звеньях. В силу равенств (18.18) и (18.19) имеем

$$q' - q = 2\varepsilon\left(\frac{1}{2} - \zeta\right)n + \\ + \sqrt{\varepsilon(1-\varepsilon)(1-\zeta)n}\, y - \sqrt{\varepsilon(1-\varepsilon)\zeta n}\, y. \qquad (18.20)$$

Алгебраическая сумма последних двух слагаемых в правой части равенства (18.20) представляет собой разность двух случайных функций, распределенных по нормальному закону. Мы увидим сейчас, что эта разность снова изображает случайную функцию с нормальным законом распределения.

Рассмотрим две случайные переменные z_1 и z_2, обладающие нормальным законом распределения со средним значением, равным нулю, и средними отклонениями σ_1 и σ_2 соответственно. Поэтому

$$\left.\begin{aligned} z_1 &= \sigma_1 y, \\ z_2 &= \sigma_2 y. \end{aligned}\right\} \tag{18.21}$$

Обозначая еще через $W_1(z_1)$ функцию распределения вероятностей переменной z_1, а через $W_2(z_2)$ — функцию распределения вероятностей переменной z_2, напишем, что

$$W_1(z_1) = \frac{1}{\sigma_1\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{z_1}{\sigma_1}\right)^2},$$

$$W_2(z_2) = \frac{1}{\sigma_2\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{z_2}{\sigma_2}\right)^2}.$$

Считая теперь эти две случайные переменные *независимыми* одна от другой, выпишем вероятность одновременного нахождения переменной z_1 в промежутке между z_1 и $z_1 + dz_1$ и переменной z_2 в промежутке между z_2 и $z_2 + dz_2$:

$$W_1(z_1)\, W_2(z_2)\, dz_1\, dz_2.$$

Введем новые переменные x_1 и x_2, определенные посредством соотношений

$$x_1 = z_1 - z_2,$$

$$x_2 = z_1 + z_2,$$

$$z_1 = \frac{1}{2}(x_1 + x_2),$$

$$z_2 = \frac{1}{2}(x_2 - x_1).$$

Упомянутая вероятность одновременного нахождения z_1 и z_2 в указанных промежутках в новых переменных принимает вид

$$\frac{1}{2} W_1\left(\frac{x_1 + x_2}{2}\right) W_2\left(\frac{x_2 - x_1}{2}\right) dx_1\, dx_2.$$

Проинтегрировав эту вероятность одновременного наступления указанных событий по x_2 в пределах от $-\infty$ до ∞, получим вероятность $W(x_1)\,dx_1$, где $W(x_1)$ представляет собой функцию распределения вероятностей переменной $x_1 = z_1 - z_2$. Таким образом,

$$W(x_1) = \frac{1}{2}\int_{-\infty}^{\infty} W_1\left(\frac{x_1+x_2}{2}\right) W_2\left(\frac{x_2-x_1}{2}\right) dx_2 =$$

$$= \frac{1}{4\pi\sigma_1\sigma_2}\int_{-\infty}^{\infty} e^{-\frac{1}{2}\left[\left(\frac{x_1+x_2}{2\sigma_1}\right)^2+\left(\frac{x_2-x_1}{2\sigma_2}\right)^2\right]} dx_2 =$$

$$= \frac{1}{4\pi\sigma_1\sigma_2} e^{-\frac{1}{2}\frac{x_1^2}{\sigma_1^2+\sigma_2^2}} \int_{-\infty}^{\infty} e^{-\xi^2} \frac{d\xi}{\sqrt{(1/8\sigma_1^2)+(1/8\sigma_2^2)}} =$$

$$= \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{\sigma_1^2+\sigma_2^2}} e^{-\frac{1}{2}\left(\frac{x_1}{\sqrt{\sigma_1^2+\sigma_2^2}}\right)^2}$$

Поэтому мы можем написать, что

$$z_1 - z_2 = \sqrt{\sigma_1^2+\sigma_2^2}\,y. \qquad (18.22)$$

Выполняя подобным же образом интегрирование по x_1, можно показать, что

$$z_1 + z_2 = \sqrt{\sigma_1^2+\sigma_2^2}\,y. \qquad (18.23)$$

Отсюда следует, что разность или сумма двух независимых случайных функций, обладающих нормальным законом распределения вероятностей, опять представляет собой случайную функцию с нормальным законом распределения вероятностей, причем квадрат ее среднего отклонения равен сумме квадратов средних отклонений исходных случайных функций. Это свойство сохранения случайных функций с нормальным законом распределения и нулевыми средними значениями по отношению к сложению и вычитанию является закономерным, так как такие функции обладают равными вероятностями при положительных или отрицательных значениях.

С помощью равенства (18.22) мы можем записать результат, выражаемый соотношением (18.20), в виде

$$q' - q = 2\varepsilon\left(\frac{1}{2} - \zeta\right)n + \sqrt{\varepsilon(1-\varepsilon)n}\,y.$$

Пусть величина $(q' - q)/n = \Delta\zeta$ представляет собой поправку к относительному числу ζ на несовершенство звеньев Шеффера; в этом случае будем иметь

$$\Delta\zeta = 2\varepsilon\left(\frac{1}{2} - \zeta\right) + \sqrt{\frac{\varepsilon(1-\varepsilon)}{n}}\,y. \qquad (18.24)$$

Теперь мы можем сочетать соотношения (18.16) и (18.24) и записать выражение скорректированного относительного числа ζ' возбужденных линий выхода в виде

$$\zeta' = \zeta + \Delta\zeta = \zeta + 2\varepsilon\left(\frac{1}{2} - \zeta\right) + \sqrt{\frac{\varepsilon(1-\varepsilon)}{n}}\,y =$$
$$= (1 - \xi\eta) + 2\varepsilon\left(\xi\eta - \frac{1}{2}\right) +$$
$$+ (1 - 2\varepsilon)\sqrt{\frac{\xi(1-\xi)\eta(1-\eta)}{n}}\,y + \sqrt{\frac{\varepsilon(1-\varepsilon)}{n}}\,y. \qquad (18.25)$$

Сумма последних двух слагаемых в правой части равенства (18.25) является суммой двух независимых случайных переменных с нормальным распределением, и, следовательно, мы можем воспользоваться соотношением (18.23). Поэтому в окончательном виде, применяя символ ζ вместо ζ', получим с помощью равенства (18.24)

$$\zeta = (1 - \xi\eta) + 2\varepsilon\left(\xi\eta - \frac{1}{2}\right) +$$
$$+ \sqrt{\frac{(1-2\varepsilon)^2\,\xi(1-\xi)\,\eta(1-\eta) + \varepsilon(1-\varepsilon)}{n}}\,y, \qquad (18.26)$$

где y — случайная переменная с нормальным законом распределения, среднее значение которой равно нулю, а среднее квадратичное отклонение — единице. Соотношение (18.26) характеризует качество срабатывания исполнительного устройства нашей системы, построенной по методу сложных сочетаний и определяемой относительными числами ξ и η возбужденных входов и относительным

числом ζ возбужденных выходов при вероятности ε отказа в срабатывании отдельных звеньев Шеффера.

18.5. Ошибки в системах, построенных по методу сложных сочетаний. Подсчитав вероятность, характеризующую качество действия исполнительного устройства нашей системы, образованной по методу сложных сочетаний, мы очень легко выполним оставшуюся часть расчета. Каждый агрегат восстанавливающего устройства, показанный на фиг. 152, в действительности равносилен исполнительному устройству. На первой ступени восстанавливающего устройства входами служат линии, образовавшиеся в результате расщепления выходных линий исполнительного устройства. Таким образом, вместо двух различных относительных чисел ξ и η мы встречаемся здесь с одним и тем же относительным числом ζ. Тогда, обозначив через μ относительное число возбужденных выходов на первой ступени, получим, в силу (18.26),

$$\mu = (1-\zeta^2) + 2\varepsilon\left(\zeta^2 - \frac{1}{2}\right) + \sqrt{\frac{(1-2\varepsilon)^2\zeta^2(1-\zeta)^2 + \varepsilon(1-\varepsilon)}{n}}\,y. \tag{18.27}$$

Подобным же образом, обозначив через ν относительное число возбужденных выходов на второй ступени восстанавливающего устройства, получим

$$\nu = (1-\mu^2) + 2\varepsilon\left(\mu^2 - \frac{1}{2}\right) + \sqrt{\frac{(1-2\varepsilon)^2\mu^2(1-\mu)^2 + \varepsilon(1-\varepsilon)}{n}}\,y. \tag{18.28}$$

Первые члены в правых частях равенств (18.27) и (18.28) совпадают с первым членом в правой части (18.1). Дополнительные члены возникают за счет несовершенства самих звеньев и за счет статистического распределения ошибок.

При любых заданных ξ, η, ε и n соотношения (18.26) — (18.28) позволяют нам подсчитать функцию распределения ν, характеризующую относительное число возбужденных выходных линий полной системы соединений Шеффера. Мы можем пояснить это обстоятельство, обращаясь к виду функции распределения. Так, например,

соотношение (18.26) равносильно соотношению

$$W(\zeta;\ \xi,\ \eta;\ n)=$$

$$=\frac{\exp\left\{-\frac{1}{2}\left[\frac{\zeta-\left[(1-\xi\eta)+2\varepsilon\left(\xi\eta-\frac{1}{2}\right)\right]}{\sqrt{\frac{(1-2\varepsilon)^2\,\xi(1-\xi)\,\eta(1-\eta)+\varepsilon(1-\varepsilon)}{n}}}\right]^2\right\}}{\sqrt{2\pi\frac{(1-2\varepsilon)^2\,\xi(1-\xi)\,\eta(1-\eta)+\varepsilon(1-\varepsilon)}{n}}}$$

Таким образом, функция $W(\nu;\ \xi,\ \eta;\ n)$ распределения вероятностей функции ν получается путем интегрирования по ζ и μ функции одновременной вероятности по переменным ζ, μ и ν, т. е.

$$W(\nu;\ \xi,\ \eta;\ n)=\frac{1}{(2\pi)^{3/2}}\times$$

$$\times\frac{1}{\sqrt{\frac{(1-2\varepsilon)^2\,\xi(1-\xi)\,\eta(1-\eta)+\varepsilon(1-\varepsilon)}{n}}}\int_{-\infty}^{\infty}d\mu$$

$$\int_{-\infty}^{\infty}\frac{d\zeta}{\sqrt{\frac{(1-2\varepsilon)^2\,\zeta^2(1-\zeta)^2+\varepsilon(1-\varepsilon)}{n}\,\frac{(1-2\varepsilon)^2\,\mu^2(1-\mu)^2+\varepsilon(1-\varepsilon)}{n}}}\times$$

$$\times\exp\left\{-\frac{1}{2}\left[\frac{\zeta-\left\{(1-\xi\eta)+2\varepsilon\left(\xi\eta-\frac{1}{2}\right)\right\}}{\sqrt{\frac{(1-2\varepsilon)^2\,\xi(1-\xi)\,\eta(1-\eta)+\varepsilon(1-\varepsilon)}{n}}}\right]^2-\right.$$

$$-\frac{1}{2}\left[\frac{\mu-\left\{(1-\zeta^2)+2\varepsilon\left(\zeta^2-\frac{1}{2}\right)\right\}}{\sqrt{\frac{(1-2\varepsilon)^2\,\zeta^2(1-\zeta)^2+\varepsilon(1-\varepsilon)}{n}}}\right]^2-$$

$$\left.-\frac{1}{2}\left[\frac{\nu-\left\{(1-\mu^2)+2\varepsilon\left(\mu^2-\frac{1}{2}\right)\right\}}{\sqrt{\frac{(1-2\varepsilon)^2\,\mu^2(1-\mu)^2+\varepsilon(1-\varepsilon)}{n}}}\right]^2\right\}.\quad(18.29)$$

Теперь мы покажем, что при определенных условиях имеется возможность достичь путем повышения числа n

почти идеального действия системы, образованной по методу сложных сочетаний из звеньев Шеффера. Рассмотрим некоторый данный критический уровень δ. Идеальный характер действия требует, чтобы из условия $\nu \leqslant \delta$ сохранения выхода в невозбужденном состоянии следовали условия $\xi \geqslant 1-\delta$ и $\eta \geqslant 1-\delta$ возбуждения обоих входов, а из условия $\nu \geqslant 1-\delta$ следовала одна из двух пар условий: $\xi \leqslant \delta$ и $\eta \geqslant 1-\delta$ или $\xi \geqslant 1-\delta$ и $\eta \leqslant \delta$. Пусть n настолько велико, а ε настолько мало, чтобы в соотношениях (18.26) — (18.28) можно было пренебречь членами порядка ε и $1/\sqrt{n}$. Тогда

$$\zeta \approx 1-\xi\eta; \qquad \mu \approx 1-\zeta^2; \qquad \nu \approx 1-\mu^2$$

или иначе

$$\nu \approx 1-(2\xi\eta-\xi^2\eta^2)^2, \quad \text{когда} \quad n \gg 1, \quad \varepsilon \ll 1. \tag{18.30}$$

Пусть теперь $\xi = 1-\alpha$, $\eta = 1-\beta$ и $\alpha \leqslant \delta$, $\beta \leqslant \delta$, так что $\xi \geqslant 1-\delta$ и $\eta \geqslant 1-\delta$. Тогда, в силу соотношения (18.30), получим

$$\nu \approx 2(\alpha^2+\beta^2)+\ldots .$$

Отсюда следует, что $\nu = O(\delta^2)$. Подобным же образом из (18.30) вытекает, что $\nu = 1-O(\delta^2)$ при $\xi \leqslant \delta$ и $\eta \geqslant 1-\delta$ или $\xi \geqslant 1-\delta$ и $\eta \leqslant \delta$. Кроме того, (18.30) также показывает, что $\nu = 1-O(\delta^4)$ при $\xi \leqslant \delta$ и $\eta \leqslant \delta$. Поэтому при малых ε и δ и при $n \to \infty$ можно добиться идеальной надежности системы, полученной по методу сложных сочетаний из звеньев Шеффера.

Когда n велико, но не бесконечно, расчет оказывается довольно громоздким из-за необходимости оценивать интеграл в соотношении (18.29). Хотя асимптотические значения этого интеграла можно найти с помощью классических методов, мы здесь не будем заниматься соответствующим расчетом. Вместо этого мы приведем один пример, заимствованный у Неймана, где $\delta=0{,}07$, т. е. возбуждение по крайней мере 93% линий связки определяет положительную информацию; возбуждение числа линий, не превышающего самое большее 7% линий, определяет отрицательную информацию.

Нейман нашел, что для предупреждения ошибок вероятность неправильного срабатывания отдельных звеньев

Шеффера должна быть ниже 0,0107. При $\varepsilon \geqslant 0{,}0107$ нельзя добиться понижения вероятности всей системы в целом до сколь угодно малой величины путем увеличения числа n. При $\varepsilon=0{,}005$ ($^1/_2$% вероятности отказа срабатывания отдельного звена) Нейман получил числовые результаты, приведенные в табл. 5. Из этой таблицы видно, что если связка насчитывает 1000 линий, то степень надежности системы довольно слабая. В самом деле, она меньше исходной надежности, характеризуемой вероятностью 1% от ε выхода из строя. Но при увеличении числа n в 25 раз достигается чрезвычайно высокая степень надежности.

Методику синтеза сложных сочетаний, рассмотренную в предыдущем разделе, можно без изменений применять также и к системам, которые с самого начала строились из звеньев Шеффера. Для этого мы заменяем каждый элемент Шеффера в исходной системе системой $3n$ звеньев Шеффера, причем каждый из них снабжаем своим исполнительным устройством и восстанавливающим устройством. Ошибку во всей системе можно подсчитать, исходя из ошибок в отдельных системах Шеффера и применяя способ, рассмотренный выше. Практически это вычисление весьма утомительно. Однако для оценки порядка величины числа n, требуемого для обеспечения заданной степени устойчивости, мы можем считать всю систему в целом равносильной одному звену Шеффера и непосредственно применять выводы для этого звена ко всей системе, что и будет сделано в следующем пункте.

Таблица 5

Число линий, n	Вероятность неправильного срабатывания
1 000	$2{,}7\times10^{-2}$
2 000	$2{,}6\times10^{-3}$
3 000	$2{,}5\times10^{-4}$
5 000	$0{,}4\times10^{-6}$
10 000	$1{,}6\times10^{-10}$
20 000	$2{,}8\times10^{-19}$
25 000	$1{,}2\times10^{-23}$

186. Примеры. Чтобы получить представление о величине объема требуемой связки, рассмотрим счетную машину, содержащую 2500 электронных ламп. Допустим, что в среднем каждая лампа возбуждается один раз каждые 5 микросекунд. Потребуем, чтобы машина в среднем работала в течение 8 часов до того, как она допустит одну ошибку. В течение этого времени число возбуждений одной лампы равно

$$\frac{1}{5}\times 8\times 3600\times 10^6 = 5{,}76\times 10^9.$$

Будем рассматривать каждую электронную лампу как звено Шеффера. Тогда, считая каждую лампу независимой, найдем, что требуемая верхняя вероятность неправильного срабатывания равна $1/(5{,}76\cdot 10^9)$.

Но в системе насчитывается 2500 связанных между собой ламп. Ошибка в срабатывании лишь одной из 2500 ламп означает, что допущена ошибка во всей системе. Поэтому, рассматривая каждую лампу, как один составной элемент системы, мы увидим, что заданная вероятность неправильного срабатывания должна составлять всего лишь 1/2500 от ранее найденного значения, т. е. равняться $1/(2500\cdot 5{,}76\cdot 10^9)=7\cdot 10^{-14}$. Мы видим, что окончательная вероятность неправильного срабатывания имеет ту же величину, которую мы получили бы, рассматривая всю систему, состоящую из 2500 ламп, как единое суммарное звено Шеффера. Эта возможность значительно облегчает подсчет требуемого числа n линий в системе, образованной по методу сложных сочетаний.

Если принять значения критического уровня δ и вероятности ε повреждения лампы, указанные в табл. 5, то для снижения вероятности неправильного действия устройства до нужного значения потребуется в соответствии с таблицей, чтобы было $n=14\,000$. Поэтому, чтобы придать машине заданную степень надежности, потребуется усложнить систему в 14 000 раз. Это означает, что каждая отдельная лампа в машине подлежит замене на узел, состоящий из $3n=3\times 14\,000=42\,000$ ламп. При этом машина, первоначально содержащая 2500 ламп, превратится в гигантское устройство из 105 000 000 ламп. Очевидно, что такой метод практически непригоден.

Обратимся теперь к другому примеру и рассмотрим правдоподобное количественное описание устройства нервной системы человека. Число нейронов, образующих эту систему, обычно определяется как 10^{10}. Но, учитывая наличие синаптических концевых узелков и других возможных автономных «подузлов», мы обнаружим, что это число, безусловно, оказывается весьма заниженным. В действительности оно в несколько сот раз больше. Примем число основных элементов равным 10^{13}. Нейрон может быть возбужден до 200 раз в секунду. Но среднее число возбуждений нейрона должно быть гораздо меньшим; примем его равным 10 возбуждениям в секунду. Допустим далее, что всякая ошибка в действии нашей нервной системы представляет собой серьезное явление и не должна иметь места на протяжении промежутка времени, соответствующего человеческой жизни. Примем интервал времени, на протяжении которого такие ошибки отсутствуют, равным 10 000 лет. В течение этого времени полное число возбуждений нервной системы, состоящей из 10^{13} элементов, составит

$$10^{13} \times 10\,000 \times 31\,536\,000 \times 10 = 3{,}2 \times 10^{25}.$$

Таким образом, вероятность неправильного срабатывания окажется равной

$$\frac{1}{3{,}2 \times 10^{25}} = 3{,}2 \times 10^{-26}.$$

Примем вновь допущение о том, что свойства основных элементов нервной системы определяются данными табл. 5. Тогда, экстраполируя значения этой таблицы, получим, что $n = 28\,000$.

Однако в этот расчет надо внести поправку: если нервная система человека действительно образована 28 000-кратным применением метода сложных сочетаний, то число основных элементов исходной системы уже не равно 10^{13}, как мы предположили выше, а в $1/(3 \times 28\,000)$ раз меньше 10^{13}. Тогда вероятность неправильного действия увеличится в $3 \times 28\,000$ раз. С учетом этой поправки вероятность неправильного действия будет равна $2{,}7 \times 10^{-21}$. Тогда с помощью табл. 5 найдем, что $n = 22\,000$. Дальнейшее построение итераций уже не приведет к заметному изменению этого числа.

Эти примеры показывают, что, хотя наш метод предупреждения ошибок с помощью сложных сочетаний вполне оправдывается в применении к микрокомпонентам нервной системы, в условиях современной техники, он оказывается, тем не менее, непрактичным в области систем автоматического регулирования, существующих на практике. Одно очевидное направление будущего развития состоит в уменьшении как размеров элементов систем, так и их мощности. С этой точки зрения полупроводниковый прибор представляет собой большое усовершенствование по сравнению с электронной лампой. Поэтому метод сложных сочетаний может еще оказаться полезным в будущем. Другое направление исследований заключается в более глубоком изучении процесса предупреждения ошибок. Метод построения нашей основной системы с помощью исполнительного и восстанавливающего элементов для каждого звена Шеффера не является, в конечном счете, единственно возможным. Хорошо, что уже эта попытка оказалась успешной в отношении выявления возможности повышения надежности действия системы. Вероятно, существуют и другие методы построения системы дублированных компонентов, доставляющие ту же степень надежности при меньшем числе таких компонентов. Другими словами, в технике предупреждения ошибок систем автоматического регулирования заложено лишь начало. Еще не существует решения этой задачи, пригодного для использования в технике автоматического регулирования.

ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА

1. Андронов А. А., Хайкин С. Э., Теория колебаний, М.—Л., 1937.
2. Ляпунов А. М., Общая задача об устойчивости движения, М.—Л., 1945, 2 изд.
3. Четаев Н. Г., Устойчивость движения, М.—Л., 1955.
4. Дубошин Г. Н., Основы теории устойчивости движения, изд-во МГУ, 1952.
5. Труды II-го Всесоюзного Совещания по теории и методам автоматического регулирования, т. I, II, III, АН СССР, 1955.
6. Лурье А. И., Некоторые нелинейные задачи теории автоматического регулирования, М.—Л., 1951.
7. Летов А. М., Устойчивость нелинейных регулируемых систем, М.—Л., 1955.
8. Немыцкий В. В., Степанов В. В., Качественная теория дифференциальных уравнений, М.—Л., 1941.
9. «Основы автоматического регулирования», сб. под ред. В. В. Солодовникова (там же приведена богатая библиография), М., 1954.
10. Конторович М. И., Операционное исчисление и нестационарные явления в электрических цепях, М.—Л., 1949.
11. Лурье А. И., Операционное исчисление, М.—Л., 1950.
12. Диткин В. А., Кузнецов П. И., Справочник по операционному исчислению, М.—Л., 1951.
13. Булгаков Б. В., Колебания, М.—Л., 1954.
14. Лаврентьев М. А., Шабат Б. В., Методы теории функций комплексного переменного, М.—Л., 1951.
15. Гарднер и Бернс, Переходные процессы в линейных системах, М., 1950.
16. Фельдбаум А. А., Электрические системы автоматического регулирования, М., 1954.
17. Маркушевич А. И., Теория аналитических функций, М.—Л., 1950.
18. Привалов И. И., Введение в теорию функций комплексного переменного, М.—Л., 1953.
19. Фукс Б. А., Шабат Б. В., Функции комплексного переменного, М.—Л., 1949.

20. Смирнов В. И., Курс высшей математики, М., 1951, т. III.
21. Джеймс Х., Никольс Н., Филлипс Р., Теория следящих систем, М., 1951.
22. Воронов А. А., Элементы теории автоматического регулирования, М., 1950, 1953.
23. Солодовников В. В., Топчеев Ю. И., Крутикова Г. В., Частотный метод построения переходных процессов, М.—Л., 1955.
24. Остославский И. В., Калачев Г. С., Продольная устойчивость и управляемость самолета, М., 1951.
25. Закс Н. А., Основы экспериментальной аэродинамики, М., 1953.
26. Сикорд Ч., Вопросы ракетной техники, № 6, 78—96 (1952).
27. Эндрью Дж., Поль Дж., Вопросы ракетной техники, № 4, 15—17 (1953).
28. Глаус Р., Вопросы ракетной техники, № 5, 107—126 (1953).
29. Цыпкин Я. З., Автоматика и телемеханика, № 2—3, 107—128 (1946).
30. Цыпкин Я. З., Бромберг П. В., Изв. АН СССР, № 12, 1163—1168 (1945).
31. Курош А. Г., Курс высшей алгебры, М.—Л., 1952.
32. «Корректирующие цепи в автоматике», сб. статей, М., 1954.
33. Айзерман М. А., Теория автоматического регулирования двигателей, М.—Л., 1952.
34. Крылов А. Н., Соч., АН СССР, т. V.
35. Кочин Н. Е., Розе Н. В., Теоретическая гидромеханика, М.—Л., 1938.
36. Мур Дж. Р. и др., Следящие и авторегулируемые системы, отличающиеся от обычных систем с обратной связью, Прикл. мех. и машиностр., сб. пер., № 5 (1953).
37. Вознесенский И. И., За советское энергооборудование, сб. статей, 1934.
38. Голомб М., Уздин Е., Теория многомерных сервосистем, Прикл. мех. и машиностр., сб. пер., № 2 (1953).
39. Гантмахер Р. Ф., Теория матриц, М.—Л., 1953.
40. Фрезер Р., Дункан В., Коллар А., Теория матриц, М., 1950.
41. Айзерман М. А., Введение в динамику автоматического регулирования двигателей, М., 1950.
42. Попов Е. П., Динамика систем автоматического регулирования, М.—Л., 1954.
43. Гольдфарб Л. С., Автоматика и телемеханика, № 5, 349; № 6, 413 (1948).
44. «Автоматическое регулирование», сб. статей, М., 1954.
45. Цыпкин Я. З., Автоматика и телемеханика, № 3, 189—224 (1949).
46. Айзерман М. А., Достаточное условие устойчивости одного класса динамических систем с переменными параметрами, Прикл. матем. и мех., **15**, № 3 (1951).

47. Записки семинара по теории устойчивости под рук. Н. Д. Моисеева, Труды гос. союзн. науч.-иссл. ин-та, науч.-техн. инф. отд., вып. 1(5) 1953.
48. Колмогоров А. Н., Фомин С. В., Элементы теории функций и функционального анализа, Изд-во МГУ, 1954.
49. Ведров В. С., Динамическая устойчивость самолета, М., 1938.
50. Цыпкин Я. З., Переходные и установившиеся процессы в импульсных цепях, М., 1951.
51. Бромберг П. В., Устойчивость и автоколебания импульсных систем регулирования, М., 1953.
52. Степанов В. В., Курс дифференциальных уравнений, М.—Л., 1937.
53. Стокер Дж., Нелинейные колебания, М., 1952.
54. Голубев В. В., Лекции по аналитической теории дифференциальных уравнений, М.—Л., 1941.
55. Льюис, Нелинейные обратные связи, Вопросы ракетной техники, № 4 (1954).
56. Боголюбов Н. Н., Митропольский Ю. А., Асимптотические методы в теории нелинейных колебаний, М.—Л., 1955.
57. Малкин И. Г., Методы Ляпунова и Пуанкаре в теории нелинейных колебаний, М.—Л., 1949.
58. Стрелков С. П., *Введение в теорию колебаний*, М.—Л., 1950.
59. Папалекси Н. Д., Собр. соч., АН СССР, 1948.
60. Горелик Г. С., Колебания и волны, М., 1950.
61. Андронов А. А., Витт А., Журн. техн. физ., IV, вып. 1, 122—157 (1934).
62. Теодорчик К. Ф., Автоколебательные системы, М.—Л., 1948.
63. Булгаков Б. В., Прикладная теория гироскопов, М., 1954, 2 изд.
64. Мак-Лахлан Н. В., Теория и приложения функций Матьё, М., 1953.
65. Задэ Л. А., Прикладная механика, сб. переводов, № 1, 1952.
66. Чаплыгин С. А., Соч., АН СССР, 1950.
67. Голубев В. В., Лекции по теории крыла в плоскопараллельном потоке, М.—Л., 1949.
68. Космодемьянский А. А., Учен. зап. МГУ, вып. 154, 73—80 (1951).
69. Гантмахер Ф. Р., Левин М. Л., Прикл. матем. и мех., т. X, 301—312 (1947).
70. Эльсгольц Л. Э., Обыкновенные дифференциальные уравнения, М.—Л., 1950.
71. Моисеев Н. Д., Записки семинара по теории устойчивости движения, ВВА им. Н. Е. Жуковского, вып. 3, 41—69, 71—85 (1948).
72. Фельдбаум А. А., Автоматика и телемеханика, IX, вып. 1, 3—19 (1948).

73. Уланов Г. М., Автоматика и телемеханика, IX, вып. 3, 168—175 (1948).
74. Красовский А. А., Труды ВВА им. Н. Е. Жуковского, вып. 281 (1948).
75. Лаврентьев М. А., Люстерник Л. А., Курс вариационного исчисления, М.—Л., 1938.
76. Фельдбаум А. А., Автоматика и телемеханика, № 4 (1949).
77. Фельдбаум А. А., Автоматика и телемеханика, № 6 (1953).
78. Фельдбаум А. А., Автоматика и телемеханика, № 2 (1955).
79. Лернер А. Я., Автоматика и телемеханика, № 2 и № 4 (1952).

УКАЗАТЕЛЬ

ОГЛАВЛЕНИЕ

［Ｇｅｎｅｒａｌ Ｉｎｆｏｒｍａｔｉｏｎ］
书名＝ТЕХНИЧЕСКАЯ КИБЕРНЕТИКА （工程控制论）
作者＝
页数＝４６２
ＳＳ号＝０
出版日期＝
Ｖｓｓ号＝５４３３９８９５